WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



(51) International Patent Classification 7:		(11) International Publication Number:	WO 00/37105
A61K 39/09, 39/40, A61P 31/04 // C07K 14/315	A2	(43) International Publication Date:	29 June 2000 (29.06.00)
 (21) International Application Number: PCT/US (22) International Filing Date: 21 December 1999 ((30) Priority Data: 60/113,048 21 December 1998 (21.12.9) (71) Applicant: MEDIMMUNE, INC. [US/US]; 35 West Mill Road, Gaithersburg, MD 20878 (US). (72) Inventors: JOHNSON, Leslie, S.; 13545 Ambassad Germantown, MD 20874 (US). KOENIG, Sco Ralston Road, Rockville, MD 20852 (US). Al John, E.; 20822 Shamrock Glen Circle, Germant 20874 (US). (74) Agents: GRANT, Alan, J. et al.; Carella, Byrne, Bain, Cecchi, Stewart & Olstein, 6 Becker Farm Road, NJ 07068 (US). 	21.12.9 8) U t Watki dor Driv tt; 107 DAMO own, M	BY, CA, CH, CN, CR, CU, CZ, GB, GD, GE, GH, GM, HU, IE KR, KZ, LC, LK, LR, LS, LT, MK, MN, MW, MX, NO, NZ, SG, SI, SK, SL, TJ, TM, TR, T YU, ZA, ZW, ARIPO patent (G SL, SZ, TZ, UG, ZW), Eurasian KZ, MD, RU, TJ, TM), Europea DE, DK, ES, FI, FR, GB, GR, SE), OAPI patent (BF, BJ, CF, C ML, MR, NE, SN, TD, TG). Published Without international search re, upon receipt of that report.	DE, DK, DM, EE, ES, FI, D, IL, IS, JP, KE, KG, KP, LU, LV, MA, MD, MG, PL, PT, RO, RU, SD, SE TT, TZ, UA, UG, UZ, VN CH, GM, KE, LS, MW, SD patent (AM, AZ, BY, KG un patent (AT, BE, CH, CY IE, IT, LU, MC, NL, PT CG, CI, CM, GA, GN, GW

(54) Title: STREPTOCOCCUS PNEUMONIAE PROTEINS AND IMMUNOGENIC FRAGMENTS FOR VACCINES

(57) Abstract

A vaccine composition is disclosed that comprises polypeptides and fragments of polypeptides containing histidine triad residues or coiled-coil regions, some of which polypeptides or fragments lie between 80 and 680 residues in length. Also disclosed are processes for preventing infection caused by S. pneumoniae comprising administering of vaccine compositions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HŲ	Hungary	ML	Mali	TT	Trinidad and Tobago
ВJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	us	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Vict Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

STREPTOCOCCUS PNEUMONIAE PROTEINS AND IMMUNOGENIC FRAGMENTS FOR VACCINES

5

۲

This application is based on U.S. Provisional Application No. 60/113,048, filed 21 December 1998, which is hereby incorporated in its entirety.

10

15

FIELD OF THE INVENTION

This invention relates generally to the field of bacterial antigens and their use, for example, as immunogenic agents in humans and animals to stimulate an immune response. More specifically, it relates to the vaccination of mammalian species with a polypeptide comprising at least one conserved histidine triad residue (HxxHxH) and at least one helix-forming polypeptide obtained from *Streptococcus pneumoniae* as a mechanism for stimulating production of antibodies that protect the vaccine recipient against infection by a wide range of serotypes of pathogenic *S. pneumoniae*. Further, the invention relates to antibodies against such polypeptides useful in diagnosis and passive immune therapy with respect to diagnosing and treating such pneumococcal infections.

25

20

In a particular aspect, the present invention relates to the prevention and treatment of pneumococcal infections such as infections of the middle ear, nasopharynx, lung and bronchial areas, blood, CSF, and the like, that are caused by pneumococcal bacteria.

BACKGROUND OF THE INVENTION

Streptococcus pneumoniae is a gram positive bacteria which is a major causative agent in invasive infections in animals and humans, such as sepsis, meningitis, otitis media and lobar pneumonia (Tuomanen et al. New Engl. J. Med. 322:1280-1284 (1995)). As part of the infective process, pneumococci readily bind to non-inflamed human epithelial cells of the upper and lower respiratory tract by binding to eukaryotic carbohydrates in a lectin-like manner (Cundell et al., Micro. Path. 17:361-374 (1994)). Conversion to invasive pneumococcal infections for bound bacteria may involve the local generation of inflammatory factors which may activate the epithelial cells to change the number and type of receptors on their surface (Cundell et al., Nature, 377:435-438 (1995)). Apparently, one such receptor, platelet activating factor (PAF) is engaged by the pneumococcal bacteria and within a very short period of time (minutes) from the appearance of PAF, pneumococci exhibit strongly enhanced adherence and invasion of tissue. Certain soluble receptor analogs have been shown to prevent the progression of pneumococcal infections (Idanpaan-Heikkila et al., J. Inf. Dis., 176:704-712 (1997)). A number of various other proteins have been suggested as being involved in the pathogenicity of S. pneumoniae. There remains a need for identifying polypeptides having epitopes in common from various strains of S. pneumoniae in order to utilize such polypeptides as vaccines to provide protection against a wide variety of S. pneumoniae.

25

20

Ξ,

5

10

15

SUMMARY OF INVENTION

In accordance with the present invention, there is provided vaccines and

vaccine compositions that include polypeptides obtained from *S. pneumoniae* and/or variants of said polypeptides and/or active fragments of such polypeptides.

The active fragments, as hereinafter defined, include a histidine triad residue(s) and/or coiled coil regions of such polypeptides.

The term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence from an alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The percent identity is determined as follows:

Percent Identity = [1-(C/R)] 100

15

20

25

5

10

=

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of the alignment between the Compared Sequence and the Reference Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have an aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, each being a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence in which the Percent Identity as calculated above is about

equal to or greater than a specified minimum Percent Identity than the Compared Sequence has the specified minimum Percent Identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

"Isolated" in the context of the present invention with respect to polypeptides and/or polynucleotides means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living organism is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the co-existing materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

20

25

15

4

5

10

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A-1C, respectively, report the results of three experiments using different preparations of SP36. The results demonstrate that active immunization with recombinant SP36 derived from pneumococcal strain Norway serotype 4 is able to protect mice from death in a model of pneumococcal sepsis using a heterologous strain, SJ2 (serotype 6B). In each of the three experiments shown, one hundred percent of the mice immunized

with SP36 survived for the 14-day observation period following challenge with approximately 500 cfu of pneumococci, while eighty to one hundred percent of sham-immunized mice (injected with PBS and adjuvant) died during the same period.

5

10

15

20

25

7

Figures 2A-2B show that passive administration of rabbit antiserum raised against Sp36 derived from Norway type 4 was able to protect mice in the pneumococcal sepsis model using two heterologous strains. Figure 2A shows that one hundred percent of the mice immunized with the SP36 antiserum survived the 21-day observation period after challenge with 172 CFU of strain SJ2 (serotype 6B). Eighty percent of the mice immunized with a control serum (rabbit anti-FimC) died by day 8, and ninety percent died by day 12. Figure 2B shows that 90 percent of the mice immunized with the Sp36 antiserum survived the 8-day observation after challenge with 862 CFU of strain EF6796 (serotype 6A). Ninety percent of the mice immunized with a control serum (collected before immunization) died by day 5.

Figure 3 is a western blot demonstrating the ability of antisera raised against recombinant Sp36 derived from strain Norway type 4 to react with Sp36 of heterologous strains. Total cell lysates were immunoblotted with mouse antisera to Sp36. A band representing Sp36 protein was detected in all 23 *S. pneumoniae* strains tested, which included isolates from each of the 23 pneumococcal serotypes represented in the current polysaccharide vaccine.

Figure 4 is a Southern blot showing that the Sp36 gene from Norway type 4 hybridizes with genomic DNA from 24 other pneumococcal strains, indicating the presence of similar sequences in all these strains.

Figure 5 is a western blot showing the reactivity of patient sera with Sp36. Sp36 (either full-length, panel A; N-terminal half, panel B; or C-terminal half, panel C) was electrophoresed by SDS-PAGE and transferred to nitrocellulose. Patient sera collected soon after the onset of illness (acute serum, lanes A) or eight to 30 days later (convalescent serum, lanes C) were used to probe the blots. For patients 2, 3, and 5, convalescent serum reacted more strongly with Sp36 than did the corresponding acute serum.

Figure 6 is an amino acid alignment comparison of four related pneumococcal proteins, namely Sp36A (PhtA; SEQ ID NO:8), Sp36B (PhtB; SEQ ID NO:10), Sp36D (PhtD; SEQ ID NO:4), Sp36E (PhtE; SEQ ID NO:6), respectively. Dashes in a sequence indicate gaps introduced to maximize the sequence similarity. Amino acid residues that match are boxed.

15

20

25

10

Figure 7 is a nucleotide alignment comparison of four related pneumococcal genes, namely Sp36A (PhtA; SEQ ID NO:9), Sp36B (PhtB; SEQ ID NO:11), Sp36D (PhtD; SEQ ID NO:5), Sp36E (PhtE; SEQ ID NO:7), respectively. Dashes in a sequence indicate gaps introduced to maximize the sequence similarity.

Figure 8 shows the results of immunization of mice with PhtD recombinant protein, which leads to protection from lethal sepsis. C3H/HeJ (Panel A and B) or Balb/cByJ (Panel C) mice were immunized subcutaneously with PhtD protein (15 μ g in 50 μ l PBS emulsified in 50 μ l complete Freund's adjuvant (CFA)). The recombinant PhtD protein used in protection experiments consisted of 819 amino acid residues, starting with the cysteine

(residue 20). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 µg protein with incomplete Freund's adjuvant (IFA) was administered 3 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at week 7; and sera from each group was pooled for analysis of anti-PhtD antibody by ELISA. Mice were challenged at week 8 by an intraperitonial (i.p.) injection of approximately 550 CFU S. pneumoniae strain SJ2, serotype 6B (Panel A), 850 CFU of strain EF6796, serotype 6A (Panel B) or 450 CFU of strain EF5668, serotype 4 (Panel C). In preliminary experiments, the LD₅₀ for strain SJ2 and EF6796 were determined to be approximately 10 CFU for both strains. The LD_{50} for strain EF5668 was determined to be < 5 CFU. Survival was determined in all groups over the course of 15 days following challenge. Data are presented as the percent survival for a total of 10 mice per experimental group. Two-sample Log-rank test was used for statistical analysis comparing recombinant Pht immunized mice to sham-immunized mice.

SUMMARY OF THE INVENTION

20

15

5

10

In accordance with one aspect of the present invention, there is provided a vaccine, generally in the form of a composition, that includes at least one polypeptide that is at least 90% identical to (c) a polypeptide

sequence comprising amino acids 1-819 of SEQ ID NO:4 or (ii) a polypeptide sequence comprising amino acids 1-460 of SEQ ID NO:6 or an active fragment of the foregoing.

In accordance with another aspect of the present invention, there is provided a vaccine, generally in the form of a composition, that includes an active fragment of a polypeptide that is at least 90% identical to (i) a polypeptide comprising amino acids 1-800 of SEQ ID NO:8 or (ii) a polypeptide comprising amino acids 1-800 of SEQ ID NO:10.

10

5

The term "active fragment" means a fragment that includes one or more histidine triad residues and/or one or more coiled coil regions. A "histidine triad residue" is the portion of the polypeptide that has the sequence HxxHxH wherein H is histidine and x is an amino acid other than histidine

15

A coiled coil region is the region predicted by "Coils" algorithm: Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting Coiled Coils from Protein Sequences, *Science* **252**:1162-1164.

20

In accordance with one embodiment, the active fragment includes both one or more histidine triad residues and at least one coiled coil region of the applicable polypeptide sequence. In accordance with another embodiment, the active fragment includes at least two histidine triad residues.

25

In another embodiment, the active fragment that includes at least one histidine triad residue or at least one coiled-coil region of the applicable polypeptide includes at least about ten percent of the applicable polypeptide and no more than about 85% of the applicable polypeptide.

The polypeptide of SEQ ID NO:4 includes five histidine triad residues, as follows:

amino acids 64-69; 188-193; 296-301; 541-546; and 625-630.

5

The polypeptide of SEQ ID NO:6 includes five histidine triad residues, as follows:

amino acids 63-68; 185-190; 289-294, 376-381; and 441-446.

10

15

20

25

In addition, the polypeptide of SEQ ID NO:4 includes two coiled-coil regions (amino acids 120-140 and amino acids 750-772) and the polypeptide of SEQ ID NO:6 includes one coiled-coil region (amino acids 119-152).

The polypeptide of SEQ ID NO: 8 includes the following regions:

HxxHxH: amino acids 63-68, 189-194, 309-314, 550-555, 634-639. Coiled-coils: amino acids 118-145, 406-434, 462-493, 724-751.

In accordance with a further aspect of the invention, a vaccine of the type hereinabove described is administered for the purpose of preventing or treating infection caused by *S. pneumoniae*.

A vaccine, or vaccine composition, in accordance with the present invention may include one or more of the hereinabove described polypeptides or active fragments thereof. When employing more than one polypeptide or active fragment, such two or more polypeptides and/or active fragments may be used as a physical mixture or as a fusion of two or more polypeptides or active fragments. The fusion fragment or fusion polypeptide may be produced,

for example, by recombinant techniques or by the use of appropriate linkers for fusing previously prepared polypeptides or active fragments.

In an embodiment of the invention, there is provided (a) a polypeptide that is at least 95% identical or at least 97% identical or 100% identical to (i) a polypeptide sequence comprising amino acids 1 to 819 of SEQ ID NO:4 or (ii) a polypeptide sequence comprising amino acids 1-460 of SEQ ID NO:6; or (b) an active fragment of the polypeptide of (a).

In the case where the polypeptide is a variant of the polypeptide comprising the mature polypeptide of SEQ ID NO:4 or SEQ ID NO:6, or any of the active fragments of the invention, the variation in the polypeptide or fragment is generally in a portion thereof other than the histidine triad residues and the coiled-coil region, although variations in one or more of these regions may be made.

In many cases, the variation in the polypeptide or active fragment is a conservative amino acid substitution, although other substitutions are within the scope of the invention.

20

5

10

15

In accordance with the present invention, a polypeptide variant includes variants in which one or more amino acids are substituted and/or deleted and/or inserted.

25

In another aspect, the invention relates to passive immunity vaccines formulated from antibodies against a polypeptide or active fragment of a polypeptide of the present invention. Such passive immunity vaccines can be utilized to prevent and/or treat pneumococcal infections in patients. In this manner, according to a further aspect of the invention, a vaccine can be

produced from a synthetic or recombinant polypeptide of the present invention or an antibody against such polypeptide.

In still another aspect the present invention relates to a method of using one or more antibodies (monoclonal, polyclonal or sera) to the polypeptides of the invention as described above for the prophylaxis and/or treatment of diseases that are caused by pneumococcal bacteria. In particular, the invention relates to a method for the prophylaxis and/or treatment of infectious diseases that are caused by *S. pneumoniae*. In a still further preferred aspect, the invention relates to a method for the prophylaxis and/or treatment of otitis media, nasopharyngeal, bronchial infections, and the like in humans by utilizing a vaccine of the present invention.

5

10

15

20

25

Generally, vaccines are prepared as injectables, in the form of aqueous solutions or suspensions. Vaccines in an oil base are also well known such as for inhaling. Solid forms which are dissolved or suspended prior to use may also be formulated. Pharmaceutical carriers are generally added that are compatible with the active ingredients and acceptable for pharmaceutical use. Examples of such carriers include, but are not limited to, water, saline solutions, dextrose, or glycerol. Combinations of carriers may also be used.

Vaccine compositions may further incorporate additional substances to stabilize pH, or to function as adjuvants, wetting agents. or emulsifying agents, which can serve to improve the effectiveness of the vaccine.

Vaccines are generally formulated for parental administration and are injected either subcutaneously or intramuscularly. Such vaccines can also be formulated as suppositories or for oral administration, using methods known in the art.

The amount of vaccine sufficient to confer immunity to pathogenic bacteria is determined by methods well known to those skilled in the art. This quantity will be determined based upon the characteristics of the vaccine recipient and the level of immunity required. Typically, the amount of vaccine to be administered will be determined based upon the judgment of a skilled physician. Where vaccines are administered by subcutaneous or intramuscular injection, a range of 50 to 500 µg purified protein may be given.

5

20

25

The present invention is also directed to a vaccine in which a polypeptide or active fragment of the present invention is delivered or administered in the form of a polynucleotide encoding the polypeptide or active fragment, whereby the polypeptide or active fragment is produced *in vivo*. The polynucleotide may be included in a suitable expression vector and combined with a pharmaceutically acceptable carrier.

In addition, the polypeptides of the present invention can be used as immunogens to stimulate the production of antibodies for use in passive immunotherapy, for use as diagnostic reagents, and for use as reagents in other processes such as affinity chromatography.

In another aspect the present invention provides polynucleotides which encode the hereinabove described polypeptides and active fragments of the invention. The polynucleotide of the present invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand.

In accordance with another aspect of the present invention, there is

provided

5

10

15

20

25

(A) an isolated polynucleotide that is at least 90% identical to a polynucleotide sequence encoding (i) a polypeptide comprising amino acids 1-819 of SEQ ID NO:4 or (ii) a polypeptide comprising amino acids 1-460 of SEQ ID NO:6, or

- (B) a fragment of the polynucleotide of (A) that encodes an active polypeptide fragment or
- (C) a polynucleotide that is at least 90% identical to a polynucleotide sequence encoding an active fragment of (i) a polypeptide comprising amino acids 1 to 800 of SEQ ID NO:8 or (ii) a polypeptide comprising amino acids 1 to 800 of SEQ ID NO:10.

In specific embodiments, the polynucleotide is at least 95% identical, preferably at least 97% identical, and even 100% identical to such polynucleotide sequence.

The term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of polynucleotides. The variants of the polynucleotides may be a naturally occurring allelic variant of the polynucleotides or a non-naturally occurring variant of the polynucleotides. The variants include variants in which one or more bases are substituted, deleted or inserted. Complements to such coding polynucleotides may be utilized to isolate polynucleotides encoding the same or similar polypeptides. In particular, such procedures are useful to obtain native immunogenic portions of polypeptides from different serotypes of *S. pneumoniae*, which is especially

useful in the production of "chain" polypeptide vaccines containing multiple immunogenic segments.

SEQ ID NO:5 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:4 and SEQ ID NO:7 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:6. SEQ ID NO:9 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:8, and SEQ ID NO:11 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:10. As a result of the known degeneracy of the genetic code, other polynucleotides that encode the polypeptides of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10 should be apparent to those skilled in the art from the teachings herein.

5

10

15

20

25

The polynucleotides encoding the immunogenic polypeptides described above may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptides of the present invention. The marker sequence may be, for example, a hexa-histidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptides fused to the marker in the case of a bacterial host, or, for example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., Cell, 37:767 (1984)).

The present invention also relates to vectors which include polynucleotides encoding one or more of the polypeptides of the invention, host cells which are genetically engineered with vectors of the invention and the production of such immunogenic polypeptides by recombinant techniques in an isolated and substantially immunogenically pure form.

4

5

10

15

20

25

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors comprising a polynucleotide encoding a polypeptide of the invention. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the polynucleotides which encode such polypeptides. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the <u>E. coli. lac</u> or <u>trp</u>, the phage lambda P_L promoter and other promoters known to control expression of genes in

prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

5

10

15

20

25

-;

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in <u>E.</u> coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the proteins.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as <u>E. coli</u>, <u>Streptomyces</u>, <u>Salmonella typhimurium</u>; fungal cells, such as yeast; insect cells such as <u>Drosophila S2</u> and <u>Spodoptera Sf9</u>; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter,

operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen, Inc.), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

10

15

5

٠

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacl, lacZ, T3, T7, gpt, lambda P_R, P_L and TRP. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular

Biology, (1986)).

The constructs in host cells can be used in a conventional manner to

jù,

5

10

15

20

25

produce the gene product encoded by the recombinant sequence.

Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples including the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of <u>E. coli</u> and <u>S. cerevisiae</u> TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with

translation initiation and termination sequences. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

5

10

15

20

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include <u>E. coli, Bacillus subtilis, Salmonella typhimurium</u> and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

25

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

5

10

15

20

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, a french press, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art. However, preferred are host cells which secrete the polypeptide of the invention and permit recovery of the polypeptide from the culture media.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

25

The polypeptides can be recovered and/or purified from recombinant cell cultures by well-known protein recovery and purification methods. Such methodology may include ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity

chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. In this respect, chaperones may be used in such a refolding procedure. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides that are useful as immunogens in the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated.

Procedures for the isolation of the individually expressed polypeptides may be isolated by recombinant expression/isolation methods that are well-known in the art. Typical examples for such isolation may utilize an antibody to a conserved area of the protein or to a His tag or cleavable leader or tail that is expressed as part of the protein structure.

20

25

15

5

10

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a

sequence of the present invention can be obtained by direct injection of the polypeptides into an animal.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The invention will be further described with respect to the following examples; however, the scope of the invention is not limited thereby:

20 Example 1

5

10

15

25

Active Protection with Anti-Sp36

A. Cloning, expression, and purification of SP36

The genomic DNA used as target for amplification was isolated from S. pneumoniae Norway strain (serotype 4), the same strain used for genomic sequencing. The complete sequence of the Sp36 gene (SEQ ID NO:9), and its predicted amino acid sequence (SEQ ID NO:8), are given in the Sequence Listing appended hereto. It was noted that the predicted amino acid

sequence included a hydrophobic leader sequence followed by a sequence (LSVC) similar to the consensus sequence for Type II signal peptidase (LxxC, in which both x's typically represent small amino acids). Primers (listed as SEQ ID NOS:1-3) were designed that would amplify the Sp36 gene and allow its cloning into pQE10 and expression as a histidine-tagged protein lacking the signal sequence for purification by nickel-affinity chromatography. Cloning of the fragment amplified by SEQ ID Nos 1 and 3 would result in a protein containing amino acids 2 through 800 of Sp36; cloning of the fragment amplified by SEQ ID Nos 2 and 3 would result in a protein containing amino acids 7 through 800 of Sp36 (amino acid numbers refer to SEQ ID NO:8).

B. Active Protection With Sp36 Vaccination

10

15

20

25

In each of the three experiments shown in Figures 1A-1C, C3H/HeJ mice (10/group) were immunized intraperitoneally (i.p.) with Sp36 protein (15 μg in 50 μl PBS emulsified in 50 μl complete Freund's adjuvant (CFA)). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 μg protein with incomplete Freund's adjuvant (IFA) was administered 4 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at weeks 3, 6, and 9; and sera from each group were pooled for analysis of anti-Sp36 antibody by ELISA. Mice were challenged at week 10 by an i.p. injection of approximately 500 CFU *S. pneumoniae* strain SJ2 (serotype 6B; provided by P. Flynn, St. Jude Children's Research Hospital, Memphis, TN). In preliminary experiments, the LD₅₀ of this strain was determined to be approximately 10 CFU. Mice were monitored for 14 days for survival.

The three experiments shown in Figures 1A-1C used slightly different

preparations of recombinant Sp36. The experiments shown in Figure 1A and 1B both used Sp36 containing amino acids 20-815, but different batches of protein were used in the two experiments. The experiment shown in Figure 1C used Sp36 containing amino acids 25-815.

5

10

15

20

25

¥,

In the experiment shown in Figure 1A, 9-week sera collected from the ten mice immunized with Sp36 (first batch) had an endpoint ELISA titer of 1:4,096,000. No anti-Sp36 antibody was detected in sera from shamimunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (520 cfu of pneumococci) for 14 days. Eighty percent of sham-immunized mice were dead by day 4, and the remainder survived.

In the experiment shown in Figure 1B, 9-week sera collected from the ten mice immunized with Sp36 (second batch) had an endpoint ELISA titer of >1:4,096,000. No anti-Sp36 antibody was detected in sera from shamimmunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (510 cfu of pneumococci) for 14 days. Of the sham-immunized mice, eighty percent were dead by day 4, and all died by day 9.

In the experiment shown in Figure 1C, 9-week sera collected from the ten mice immunized with Sp36 (containing amino acids 25- 815) had an endpoint ELISA titer of 1:4,096,000. No anti-Sp36 antibody was detected in sera from sham-immunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (510 cfu of pneumococci) for 14 days. Of the sham-immunized mice, ninety percent died by day 4, and all died by day 12. These data demonstrate that immunization of mice with recombinant Sp36 proteins elicits a response capable of

protecting against systemic pneumococcal infection and death. This protection was not strain-specific: the recombinant pneumococcal protein was cloned from a serotype 4 strain, while the challenge was with a heterologous strain, SJ2 (serotype 6B).

5

¥,

Example 2

Passive Protection with Anti-Sp36 Antisera

A. Generation of Rabbit Immune Sera

10

15

Following collection of preimmune serum, a New Zealand White rabbit was immunized with 250 μg of Sp36 (containing amino acids 20-815) in CFA. The rabbit was given two boosts of 125 μg Sp36 in IFA on days 29 and 50 and bled on days 39 and 60. A second rabbit was immunized with a control antigen, *E. coli* FimC.

B. Passive Protection in Mice

C3H/HeJ mice (10 mice/group) were passively immunized by two i.p. injections of 100 µl of rabbit serum. The first injection was administered twenty-four hours before challenge with 172 cfu of *S. pneumoniae* strain SJ2, and the second injection was given four hours after challenge. Figure 2 shows the survival of mice after infection with two different strains of pneumococci.

25

20

Figure 2A shows that of mice injected with 172 cfu of strain SJ2 (Figure 2A), one hundred percent of the mice immunized with rabbit immune serum raised against Sp36 protein survived the 21-day observation period. Of the mice immunized with the control serum (anti-FimC), eighty percent

died by day 8, and ninety percent died by day 12. Figure 2B shows that of mice injected with 862 cfu of strain EF6796, ninety percent of the mice immunized with rabbit immune serum raised against Sp36 protein survived the 8-day observation period. Of those given a control serum (collected from a rabbit before immunization), ninety percent died by day 8.

These data indicate that the protection against pneumococcal infection resulting from immunization with Sp36 is antibody-mediated, since mice can be protected by passive transfer of serum from a hyperimmunized rabbit. As seen in the mouse active challenge experiments described above, serum directed against recombinant Sp36 protein cloned from a serotype 4 strain was protective against challenge with heterologous strains.

Example 3

5

10

20

25

15 Conservation of Sp36 Among Strains of S. pneumoniae

A. Western blotting

The 23 pneumococcal strains used in this experiment were obtained from the American Type Culture Collection (Rockville, MD) and include one isolate each of the 23 serotypes in the multivalent pneumococcal vaccine. For total cell lysates, pneumococci were grown to mid-logarithmic phase (optical density at 620 nm, 0.4 to 0.6) in 2 ml Todd-Hewitt broth with 0.5% yeast extract (Difco, Detroit, ME) at 37°C. Bacteria were harvested by centrifugation and washed twice with water. Pellets were resuspended in 200 µl lysis buffer (0.01% sodium dodecyl sulfate, 0.15 M sodium citrate and 0.1% sodium deoxycholate) and incubated at 37°C for 30 min, then diluted in an equal volume 2x SSC (0.3 M sodium chloride, 0.03 M sodium citrate). Lysates were separated by SDS-PAGE, transferred to nitrocellulose

membranes (Bio-Rad Laboratories, Hercules, CA), and probed with antibody in a standard Western blotting procedure. Sera from ten C3H/HeJ mice immunized with Sp36 (as described in Example 1) were pooled and used at a dilution of 1:3000. Bound antibody was detected with peroxidase-conjugated sheep anti- mouse IgG using the chemiluminescence kit from Amersham, Inc. (Cambridge, MA).

The mouse anti-Sp36 sera detected two major bands with apparent molecular weights of 97 and 100 kDa in all 23 pneumococcal lysates tested (shown in Figure 3). The Sp36 signals obtained from *S. pneumoniae* serotypes 1, 5, 17F and 22F were lower, indicating either that the level of Sp36 expression is reduced in these strains, or that Sp36 in these strains is antigenically different.

These data show that Sp36 is antigenically conserved among strains of the 23 pneumococcal serotypes represented in the current polysaccharide vaccine.

20 B. Southern blotting

10

25

Genomic DNA was prepared from each of the 23 pneumococcal strains listed in the previous section and also from strain SJ2. DNA was digested with *Pvull* and *BamHI*, electrophoresed in an agarose gel and transferred to a nylon membrane. A probe was prepared by amplifying the Sp36 gene from Norway type 4 DNA (as in Example 1) and labeling the amplified fragment with fluorescein by the random-priming method, using a kit from Amersham. Hybridization, washing, and exposure of film were carried out as in the protocol supplied by Amersham. Figure 4 shows that

the Sp36 probe hybridized with DNA from each of the 24 strains studied. The lane marked "M" contained DNA from lambda phage, digested with Hindll and labeled with fluorescein, as molecular weight markers.

Example 4

5

10

15

20

Immunogenicity of Sp36 in Humans

In order to determine whether Sp36 is immunogenic during human culture-proven patients with from infection, sera pneumococcal pneumococcal bacteremia were used in Western blots containing recombinant Sp36 protein. In the experiment shown in Figure 5, sera from five patients (indicated as 1 through 5) were diluted 1:3000 and used to probe blots containing full-length Sp36, the N-terminal half of Sp36 (preceding the proline-rich region), or the C-terminal half of Sp36 (following the proline-rich region). Lanes labeled A (acute) were probed with serum collected shortly after diagnosis of pneumococcal infection; lanes C (convalescent) were probed with serum collected either one month later (patients 1, 2, and 3) or eight days after the first serum collection (patients 4 and 5). For patients 2, 3 and 5, reactivity of the convalescent serum with Sp36 was stronger that that of the corresponding acute serum. The difference between the acute and convalescent sera was particularly evident for reactivity with the C-terminal half of the protein.

In additional experiments (not shown), convalescent sera from 23 patients with pneumococcal infections were tested individually for reactivity with full-length Sp36: 20 of the 23 sera were found to bind Sp36 on a Western blot.

These experiments indicate that Sp36 is recognized by the human

immune system and suggest that antibodies able to bind the Sp36 protein may be produced during natural *S. pneumoniae* infection in humans. Since the patients were infected with a variety of pneumococcal strains, these data also support the idea that Sp36 is antigenically conserved.

5

Example 5

Table 1 provides the percent identity between the various sequences.

10

Alignment of the predicted amino acid sequences of PhtA, PhtB, PhtD, and PhtE using the MEGALIGN program of Lasergene showed strong N-terminal homology with substantial divergence of the C-termini (Figure 6). The alignment of the nucleotide sequences of the same genes is shown in Figure 7. Amino acid and nucleotide sequences were compared using the identity weighting in a Lipman-Pearson pairwise alignment, in which the number of matching residues is divided by the total of matching residues plus the number of residues in gaps. In the table below, the percent identity between each pair of sequences is shown at the intersection of the corresponding row and column.

20

25

15

Example 6

Active Protection with PhtD Vaccination.

Mice immunized with recombinant PhtD derived from strain N4 generated potent antibody titers (reciprocal endpoint titers ranging form 2,048,00 to 4,096,000). Mice immunized with PhtD were protected against death following intraperitoneal injection with either of three heterologous strains, SJ2 (serotype 6B; provided by P. Flynn, St. Jude Children's Research

Hospital, Memphis, TN), EF6796 (serotype 6A) or EF5668 (serotype 4; both strains provided by D. Briles, University of Alabama, Birmingham). In the experiment shown in Figure 8 (Panel A), all ten of the sham-immunized mice died within 10-days after challenge with virulent pneumococci (strain SJ2), while eighty percent of the PhtD-immunized mice survived the 15-day observation period. Immunization with PhtD also protected against a serotype 6A strain, EF6796 (Panel B) and a serotype 4 strain, EF5668 (Panel C). In the experiment shown in Figure 8 (Panel B), all ten of the sham-immunized mice died within 7-days after challenge with virulent pneumococci (strain EF6796), while ninety percent of the PhtD-immunized mice survived the 15-day observation period. In the experiment shown in Figure 8 (Panel C), all ten of the sham-immunized mice died within 6-days after challenge with virulent pneumoccoci (strain EF5668), while eight of nine mice immunized with PhtD survived the 15-day observation period.

Table 1. Percent Identities

	PhtA	PhtB	PhtD	PhtE
PhtA		66.4	63.9	49.5
PhtB			87.2	49.5
PhtD				49.8
PhtE				***
Percent Id	entity Between I	Nucleotide Sequ	ences PhtD	PhtE
Percent Id				PhtE 47.9
		PhtB	PhtD	
PhtA		PhtB	PhtD 59.3	47.9

WHAT IS CLAIMED IS:

5

10

15

20

25

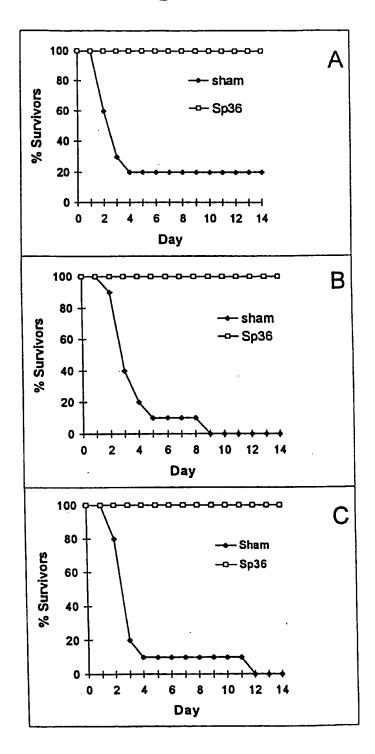
- 1. A vaccine composition comprising:
- at least one member selected from the groups consisting (a) of (i) a polypeptide comprising a polypeptide sequence that is at least 90% identical to amino acids 1-819 of SEQ ID NO:4; (ii) a polypeptide comprising a polypeptide sequence that is at least 90% identical to amino acids 1-460 of SEQ ID NO:6; (iii) a fragment of the polypeptide of (i) that includes at least one of a histidine triad residue or coiled-coil region; (iv) a fragment of the polypeptide of (ii) that includes at least one of a histidine triad residue or a coiled-coil region; (v) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:8, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids; and (vi) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:10, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids; and
 - (b) a pharmaceutically acceptable carrier.
- 2. A process for preventing infection caused by *S. pneumoniae* comprising:

administering the vaccine of claim 1.

- 3. A vaccine composition comprising:
- (a) at least one antibody against a member selected from the group consisting of (i) a polypeptide comprising a polypeptide sequence that

is at least 90% identical to amino acids 1-819 of SEQ ID NO:4; (ii) a polypeptide comprising a polypeptide sequence that is at least 90% identical to amino acids 1-460 of SEQ ID NO:6; (iii) a fragment of the polypeptide of (i) that includes at least one of histidine triad residue or coiled-coil region; (iv) a fragment of the polypeptide of (ii) that includes at least one of a histidine triad residue or a coiled-coil region; (v) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:8, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids and (vi) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:10, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids.

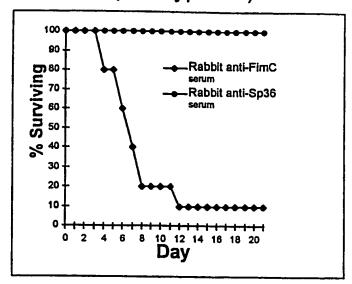
Figure 1



1/17

Figure 2

A. Strain SJ2 (serotype 6B)



B. Strain EF6796 (serotype 6A)

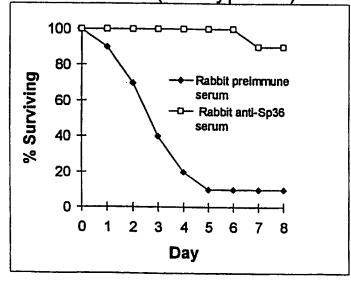


FIG.3A

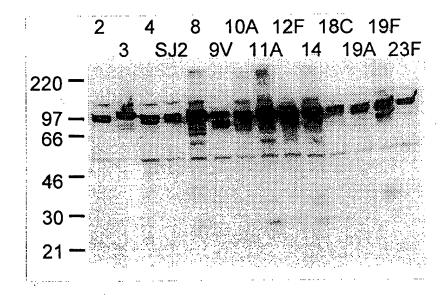
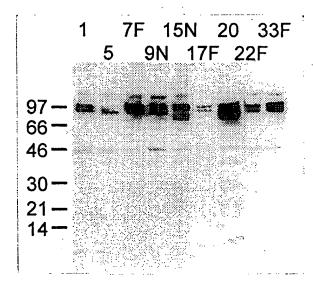
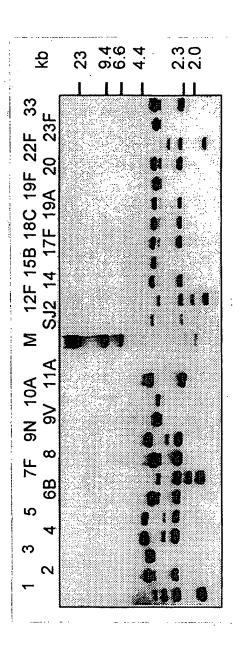


FIG.3B



i ę

F1G.4



4/17
SUBSTITUTE SHEET (RULE 26)

FIG.5A

FIG. 5B

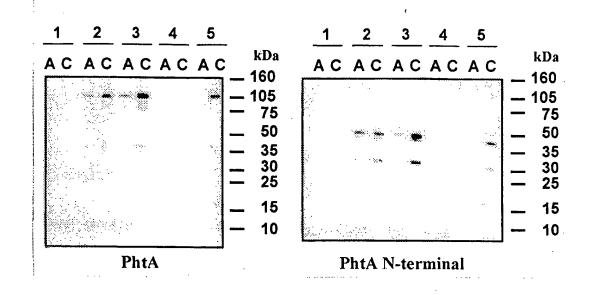
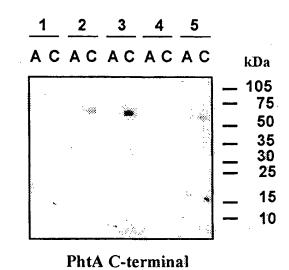


FIG.5C



5/17 SUBSTITUTE SHEET (RULE 26)

Figure 6(a)

	CSY	EL	G	R H	Q	A (G Q	X	ĸ	x	E	S 1	1 1	R V	· s	Y	I	D	0 1	0 0) A	G	0	ĸ	A	2 :	N I	٠,				· v	s	ĸ	R	E	G	ı	N	A.	E (2	H	1) 0	rit	Y
						10								2	0								30)								4	0								5	0				
1	CSY	E L	G I	R H	Q	,	G Q	٧	ĸ	K	E	S 2	1 1	2 V	3	Y	I	D	0 1	5 (λ	G	o	ĸ	λ	E	N I	. 7		. [2		S	K	R	E	C	I	N ·	۸	E	1	Pì	neD	. PR	10
1	CSY	EL	G	RY	·lo	A	G C	ما	k	ĸ	Z	s t	1	R V	<u>la</u>	Y	I	D	G	0 (À	G	Q	ĸ	λ	2	N :	6 1		י ב	E	: v	S	K	R	E	G	I	N	λ	E C	Ы	Pł	1tB	. DE	
1	CSY	EL	C)	Y	Ø	٨	- 4	T	٧	K,	E)	N	1	RV	S	Y	i	D	51	5	3	J T	0	K	1	Z :	N :	. 1		, ב	7	٦,	S	K	R	ı E	C	I	<u> </u>	λ .	E (Pł	IEA	. PR	.0
1	CAY	ᄮᄔ	N	O IH	JR	Ş	- [2	JE	N	(E)	D	N [I	1	<u> </u>	_3	- 1	٧,	<u>.</u>	<u>د</u>	א נג	Ŋ,	. 3	Q	K)	5 (N.	47	1		1 6	Y	5	10	*	E	<u> </u>	ب	αį	Α.	<u> </u>	21	Pł	ICE	. PR	O
	IVI	K I	T	D Q	a	Y	Vτ	S	ĸ	a	D.	H ?	2 1	K Y	Y	N	G	K '	V	<u> </u>	1 0		I	I	S	E	2	١ 1	. 1	1		7	N	Y	Q	Ł	ĸ	D	s	D.	t v	<u>.</u>	He	ıjo	F 1t	Y
						60								7	0								10)								9	0								10	30				
51	IVI	K I	T	0 0	G	Ÿ	V I	s	H	G	D	K 1	()	<u> </u>	Y	N	G	K '	V	PY	? 2	, <u>a</u>	7	I	5	E	E :	L 1			: E	, ,	N	Y	Q	Ł	x	D	s	D	1	7	Pì	ntD	. PR	10
51	IVI	K I	T	0 0	G	Y.	V T	5	H	Œ	D :	H 1	r 1	K Y	Y	N	G	K '	V	P 1	ſ	A	I	I	S	E :	E	6 1	. H	1	: 2	8	N	Y	0	L	K	ַם	S	ם	ľ	/	P	ntB	.pr	0
50	IVI																																													
50	IVI	K I	T	D 0	0	Y	<u>V 1</u>	3	<u> </u>	G	D	<u> </u>		<u> </u>	<u> </u>	N	<u> </u>	K	<u>V 1</u>	<u> </u>	ַ כ) <u>X</u>	L		3	<u> </u>	2	. 1			<u> </u>) [N	Y	0	Ļ	ĸ	D	٨Į	D	<u> </u>	7	Pl	htE	. 91	10
	NEV	K G	G :	r v	I	K ·	V E	a	ĸ	Y	Y.	V 3	r 1	. x	. 0	A	λ	Ħ.	A	D P	ŧ 1	/ R	T	ĸ	K	B	I	N 1	1 0	,	: 0) Z	H	3	H	N	ĸ	Z	G	0	• •	_	H	ajo	rit	Y
						110								12	0	_							נָנ	0								14	0						-		1	T 50				
101	NEI	K G	6	r v	Ī	K	V E	a	X	Y	Y.	V 1	7 1	. ,	D	Ā	A	H.	À	D 1	1	R	T	K	Z	ž	I	K I	۲ (1	C	1 2	H	S	н	N	К	G	G	G	-	-	P	htE	. PI	an out
101	NEI	K G	Ø 1	r v	I	K '	VN	a	X	Y	¥.	V٦	r 1	L	. D	A	A	H.	A I	D 1	7 1	: IR	T	x	ĸ	X	zi:	K I	t C		: 0	1	R) s	H	N	H	N	5	•	• •	-	P	htB	.pz	ro
100		K G	a :	2	ı	K '	VE	•	K	¥	Y'	V 1			: D	Y	Ÿ	H.	A I	D 3	! V	7 R	T	K	=	ĸ	I	N I		} }	: 0	3 2	H	<u>. s</u>] 0	H	R	E,	<u>c</u>	C	T	P	P	heJ	. PI	0
100	NEV	K U	<u> </u>	ı ı	Œ	K.	V D	0	X	Y	Y	<u>V 3</u>			<u>. 0</u>		^	n	^_	<u> </u>	<u> </u>	<u> </u>	<u>. T</u>	K	P	4	I					<u> </u>	. n	ΊΛ	X.	ט	N	لقا	•	-	-	-	P	uer	. PI	
	RND	X A	v :	A A	. A	R.	A Q	0	R	Y	T	7 1) (9	Y	I	?	N	λ	S 1	2	: 3	B	D	Ŧ	Q	D.	A :	1	. 3	_ 1	1	0	D	H	Y	Ħ	Y	I	P	K	N.	H) ta	r1	ĽΥ
					1	rėo	ì							17	0								1	0								19	0								2	٩o)			
	SND	-1-				-		_	_	_	_						_		_	_	_	-	-		_		_	-	-		, ,		,		_		#	Ŧ	Ŧ	5	_		٠.	hel	3 0	RO
149	SIM DI		ΑL	VJ۸	• •		~ ^		H.	¥	4.	T 1		,		-	7	N	A.	S 1	: פ		Z	Ð	Ŧ	0	D		•	٠,		•			, 4		••	•	•	•	•	"I	•			
148	RAD	иla	V	λ	A	R	A.C	a	R	¥	T	T 1	0 1	0	X	I	r	H	A.	8 1	ו כ	1		D	T	a	D	A :	r 1	t٦	, ,	7 1	1 0	3	H	Y	H	Y	I	P	K	N	₽	bti	3.p	ro
148	RAD	NA	V	N A		R R	AC		R	Y	T T	T I	D 1	0 0	X	I	r	H	A A	s i s i	D 1		×	D	T T	a a	D D	A S	r :	t 1	/ 1 / 1	7 1 7 1	1 C	3 6	H	Y	H	Y	I	P	K :	N	2	bt! bt!	1.p	ro RO
148	RAD	NA	V	N A		R R	AC		R	Y	T T	T I	D 1	0 0	X	I	r	H	A A	s i s i	D 1		×	D	T T	a a	D D	A S	r :	t 1	/ 1 / 1	7 1 7 1	1 C	3 6	H	Y	H	Y	I	P	K :	N	2	bt! bt!	1.p	ro RO
148	RAD	H A G A S N	v v	A A		R R R	3 C	0	RR	Y	T T	֓֞֞֞֜֞֜֞֝֞֜֜֝֟֝֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֓֓֓֓֓֡֝֓֓֓֡֡֝֡֓֓֓֡֓֡֓֡֓֡֡֡֝֓֡֓֡֡֝֓֡֡֡֝֓֡֡֡֝֓֡֡֡֝֡֡֡֝֡֡֡֝֓֡֡֜֝֡֜֝֡֡֜֝֡	D I	0 0	Y	V	7	и и и	A A P	8 I 8 I 8 I	0 1		1	D	T T	a G	D D 원	A S	r :	7 1	7 1	? 1 ? 1	1 0		H	Y	H H	Y	I	P P	K K	ห ห 5	5	hti hti	3.p: 1.p: 2.p:	ro RO RO
148	RAD	H A G A S N	v v	A A	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	R R R	2 J	0	RR	Y	T T	֓֞֞֞֜֞֜֞֝֞֜֜֝֟֝֓֓֓֟֝֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֓֓֡֓֡֓֡֓֓֡֓֡֓֡֓֡	D I	0 0	Y	V	7	и и и	A A P	8 I 8 I 8 I	0 1	- 0	1	D	T T	a G	D D 원	A S	r :	7 1	7 1	P 1	1 0		H	Y	H H	Y	I	P P	K K S	ห ห 5	P P	hti hti	3.p: 1.p: 2.p:	ro RO RO
148 150 146	RAD RND KVN ELS	NA SN AS	V V	A L		R R R	S (S	Y	RRL	N N	7 6	τ : τ :	מאו	22	1 Y	-	-	- - -	A P	S		- 0	23	D D S	TTRR	0 0 2	D N S	A A		3 !	7 1 7 1	21	10		1 P	X	H H Q	YY	III	PPL	K K S	N N S B 50 E	. H	hei hei hei	a.p: L.P: E.P: ori	RO RO TY
148 150 146	RAD RND ELS ELS	NA SN AS AS	V V	LA		R R A 210	2 J	Y	RRL	N X X	7 0	T X	מא	22	¥ Y	-	-	и и -	A A P	8 I		- 0	23	DDS	TTTR	2 0 0 0 P	ช ช ร ร	3 : S : S :			7 1 7 1	21	10		H	XXX	H H H	YYY	III	PPL	K K S S	N N S R 5 P E	1	hei hei hei	3.p: 1.p: 2.p: 0.p:	RO RO RO RO
148 150 146 199 198 200	RAD RND KVN ELS ELS ELS	NA SN AS AS	V V E			RRR A-11-AAA	S S S S S S S S S S S S S S S S S S S	Y	RRL	Y Y H H	7 T C C C	T X X X X	מאם	22	Y	N	777 -	- H	AAP	S I		- 0	23	D D S S S S	TTTRRRD	2 P P P P P P P P P P P P P P P P P P P	D N S S T	A S				21 21 21 41	10	N N N N N N N N N N N N N N N N N N N	H	YY	H H H O O N	YYY	RRC	PPLLT	K K S S S	N N S B S E N	1	hti hti hti hti	3.p: 1.p: 2.p: 5.p: 5.p	RO RO RO RO RO RO
148 150 146 199 198 200	RAD RND EVN ELS ELS ELS DLS	NA SN AS AS AS	V V		A A A A A A A A A A A A A A A A A A A	RR ATION	E J	Y Y Y	RRL	Y Y N N N	0000	T X X X X X X X X X X X X X X X X X X X	D 1	22	Y Y Y	- N		- K	AAP T	S S S S S S S S S S	R		23	D D S S S S S	TTTRRRDH	2 2 2 N Q	5 5 7	3 S S S S			7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1		10011	N N N N N N N N N N N N N N N N N N N	H H H H H H H H H H H H H H H H H H H	YYX	H H H O	PPP	RRC	P P L L T -	S S S S	N S E S E N N	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	hti hti hti hti hti	0.P 0.P 0.P 0.P	RO RO RO RO RO RO
148 150 146 199 198 200	RAD RND KVN ELS ELS ELS	NA SN AS AS AS	V V		A A A A A A A A A A A A A A A A A A A	RR ATION	E J	Y Y Y	RRL	Y Y N N N	0000	T X X X X X X X X X X X X X X X X X X X	D 1	22	Y Y Y	- N		- K	AAP T	S S S S S S S S S S	R		23	D D S S S S S	TTTRRRDH	2 2 2 N Q	5 5 7	3 S S S S			7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1		10011	N N N N N N N N N N N N N N N N N N N	H H H H H H H H H H H H H H H H H H H	YYX	H H H O	PPP	RRC	P P L L T -	S S S S	N S E S E N N	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	hti hti hti hti hti	0.P 0.P 0.P 0.P	RO RO RO RO RO RO
148 150 146 199 198 200	RAD RND EVN ELS ELS ELS DLS	NA SN AS AS AS	V V		A A A A A A A A A A A A A A A A A A A	RR ATION	E I E I E I E I E I E I E I E I E I E I	Y Y Y	RRL	Y Y N N N	0000	T X X X X X X X X X X X X X X X X X X X	D 1	22	200	- N		- K	AAP T	S S S S S S S S S S	R		23	D D S S S N L	TTTRRRDH	2 2 2 N Q	5 5 7	3 S S S S			7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1		10011	N N N N N N N N N N N N N N N N N N N	H H H H H H H H H H H H H H H H H H H	YYX	H H H O	PPP	RRC	P P L L T -	K K S S S S T N	N S E S E N N		hti hti hti hti hti	0.P 0.P 0.P 0.P	RO RO RO RO RO RO
148 150 146 199 198 200 196	RADRND RVN ELS ELS ELS THN	N A S A S A S A S A S A S A S A S A S A	V V V		AAA A T	RER ATION	N S S S S S S S S S S S S S S S S S S S	Y	RRL	N N N N N N N N N N N N N N N N N N N	000000	T X X X X X X X X X X X X X X X X X X X	מא	22	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	I I V	L	H H H	AAP - T E E	S I	D I		23 O O N	D D S S S S N L	TTT R RRDH S	O C P P P N O . E	D D N S S T P R	A A S S S S S S S	S !		7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		N N N N N N N N N N N N N N N N N N N	HH	YYY	HHHOOND	YYY	R R R C - O	PPLLTTI	K S 2 S S T N T	NNS BESTEEN N STOLE	T T T T T T T T T T T T T T T T T T T	hti hti hti hti hti hti	0.P 0.P 0.P 0.P 0.P	RO RO RO RO RO RO
148 150 146 199 198 200 196	RADRNO RND ELS ELS ELS DLS THN	NAS AS AS LT	V V E E E E V V V		AAAAA T	RRR ATION Y SOLYY	ASS EN RELATED HO	Y Y Y	R R R L W W L L	Y Y N N N N N N N N N N N N N N N N N N	*** a coco a co	T X X X X X X X X X X X X X X X X X X X	מא	22	200	TIV -	7 7 7 - S - L L L	H H H R R R R R R R R R R R R R R R R R	AAP - T E EE	S I			23 0 0 N	D D S S S N L	TTT R RRDH S	O C P P N O . E E	D N S S T P R R	A A S S S S H	S I			2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		N N S S S S S S S S S S S S S S S S S S	H H H P P P P P P P P P P P P P P P P P	YYY	HHH O OND P	YYY	TII R	P P L L T - I	K S 2 S S T N T T	NIS BELLENH STOLES	THE SERVICE SERVICES	hei hei hei hei hei hei hei	1.p. 1.p. 2.p. 2.p. 2.p. 2.p. 2.p. 2.p.	RO R
148 150 146 199 198 200 196	RADRNO RVN ELS ELS ELS THN NHN NHN	N A S A S A S A S A S A S A S A S A S A	V V V E E E E V V V N		AAA A TTTT	REE ATTAKA Y STYYN	ASS E A A A A A A A A A A A A A A A A A	TY THE A	R R L W W L L W W W N W W N W W N W W N W W N W W N W	N N N N N N N N N N N N N N N N N N N	*** C C C C C C C C C C C C C C C C C C	T X X X X X X X X X X X X X X X X X X X	D N	22	Y Y Y	TIV -	7 7 2 L L L L L L L L L L L L L L L L L	H H H R R R R R R R R R R R R R R R R R	A A P	S A	N I		23 O O N O O N O O O O O O O O O O O O O	DDD S O S S S H L	TTT R RRDH S	O C P PRO B	D D N S S T P R R R R	A A A S S S S H	S S S S S S S S S S S S S S S S S S S					N N S T	HH	YYY	HHHOOND	YYY P PPP A AAA	REG - O OOO	PP L LLT I	KKK S 2 SSTN T 1	NUS BEENN STORES	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	heine	D.P. D.P. D.P. D.P. D.P. D.P.	RO R
148 150 146 199 198 200 196	RADRNO RND ELS ELS ELS DLS THN	N A S A S A S A S A S A S A S A S A S A	V V V E E E E V V V N		AAA A TTTT	REE ATTAKA Y STYYN	ASS E A A A A A A A A A A A A A A A A A	TY THE A	R R L W W L L W W W W W W W W W W W W W	N N N N N N N N N N N N N N N N N N N	*** C C C C C C C C C C C C C C C C C C	T X X X X X X X X X X X X X X X X X X X	D N	22	Y Y Y	TIV -	7 7 2 L L L L L L L L L L L L L L L L L	H H H R R R R R R R R R R R R R R R R R	A A P	S A	N I		23 O O N O O N O O O O O O O O O O O O O	DDD S O S S S H L	TTT R RRDH S	O C P PRO B	D D N S S T P R R R R	A A A S S S S H	S S S S S S S S S S S S S S S S S S S					N N S T	HH	YYY	HHHOOND	YYY P PPP A AAA	REG - O OOO	PP L LLT I	KKK S 2 SSTN T 1	NUS BEENN STORES	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	heine	D.P. D.P. D.P. D.P. D.P. D.P.	RO R
148 150 146 199 198 200 196	RADRNO RVN ELS ELS ELS THN NHN NHN	N A S A S A A S A A S A A S A A S A A S A A S A A S A A S A A A S A	V V E E E E E V V N K		AAAA T TTTT	RER ATION AND YTES YYME	A S S S S S S S S S S S S S S S S S S S	A A A A A A A A A A A A A A A A A A A	R R R W W L L N	YYY N N N S A O O O K	**** C C C C C C C C C C C C C C C C C	T T K K K R K B B B B B B B B B B B B B B B	מ א ש		200	THY - L LLLL	P P - S - L LLL	H H H R R R R R R R R R R R R R R R R R	AAP - T E E C E	SAL	Y Y		23 0 0 N 2 2 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	DDD S S S S N L D LLLS	TTT R RRDH S SSA	DOC P PRIO E E E E E E	D D N S S T P R R R R R R	A A B B B B B H H H Y				2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	10 11 11 11 11 11 11 11 11 11 11 11 11 1	N N S T	HH	YYY	HHHOONDP	YYY	HHE REG O OOK	PPPLLTT	S S T N T T T T	NNS BEENN STOLESSE	2	heine	D.P D.P D.P D.P D.P D.P D.P D.P	RO R
148 150 146 199 198 200 196	RADRND RND ELS ELS ELS DLS THN NHN NHN TNT TOS	N A S A S A A A S A A A S A A A S A A A S A A A S A A S A A S A A S A A S A A A S A A S A A S A A S A A S A A S A A S A A S A A S A A S A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A S A A A A S A A A A A S A	V V E E E E V V N K V		AAA TTTT	RRR ATIONALA YTES YYME HILL	A S S S S S S S S S S S S S S S S S S S	Y Y Y H A	R R R L W W L L W W R R W W W L L	TYY N N N S A O O O O E H	*** C C C C C C C C C C C C C C C C C C	T T X X X X X X X X X X X X X X X X X X	מא שו	22 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	X X Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y Y	THV -	- S - L L L L S	H H H R R R R R R R R R R R R R R R R R	AAP - T E EEQE L	S A L	N T T T T T T T T T T T T T T T T T T T		23 0 0 N 23 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	DDD S S S N L D L L S R	TTT R RRDH S SSSA I	DOC P PRO E E E E E E E E E E E E E E E E E E E	DDNSSTPRRRRR	A A S S S S H H H Y L	K S S S S S S S S S S S S S S S S S S S		7 3 7 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	TO THE PROPERTY OF THE PROPERT	S S S S S S S S S S S S S S S S S S S	H H H H H H H H H H H H H H H H H H H	YYY A AASS D DDDDDDDDDDDDDDDDDDDDDDDDDDD	HHH O OOND P	YYY	TIT R RRG - O OOK R	PPLLTT	K K S 2 S S T N T T T T T T T T T T T T T T T T	NNS RESTRICKE STOLERS OF ST		hei hei hei hei hei hei hei hei hei	D.P. D.P. D.P. D.P. D.P. D.P. D.P. D.P.	RO R
148 150 146 199 198 200 196 238 237 250 230	RADRND RVN ELS ELS ELS DLS THN NHN TNT TOS	N A S A S A S A S A S A S A S A S A S A	V V E E E E V V N K V V	LANU LA A V T S S S S A V	AAAA TTTT P	RRR ATION AND YTES YYME HTIL	A S S S S S S S S S S S S S S S S S S S	TY YY A	R R R W W L L W W W W W W W W W W W W W	H EOOO H	*** C C C C C C C C C C C C C C C C C C	T T T T T T T T T T T T T T T T T T T	D N H H D N H	21 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	THY - LLLL M	2 7 7 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3	H H H R R R R R R R R R R R R R R R R R	AAP - T E BEOR L	S S A L L L L L L L L L L L L L L L L L	N I		23 0 0 N 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	DDD S O S S N L O LL S R O R	TTT R RRDH S SSSA I	DOC PPRO E E E DOC I	DDN S SSTP R RRR P	AAA SSSS HHHHY L	K T S S S S S S S S S S S S S S S S S S		7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1		TO COLUMN TO SERVICE STATE OF THE SERVICE STATE OF	N N S T V	HH H F F F F F F F F F F F F F F F F F	YYAAAAS	HHH O OND P	YYY	HIII R RRG - O OOK R	PPLLTT	KKK S 2 SSTN T TTTI	NUS REPUBLISHED OF STORES	TO THE SECOND SECOND	heine	D.P. D.P. D.P. D.P. D.P. D.P. D.P.	RO R
148 150 146 199 198 200 196 238 237 250 230	RADRND RVN ELS ELS ELS DLS THN NHN TNT TOS RTA	NAS AS AS LT LTTN RG	V V E E E E V V V V V V V V V V V V V V	LA LA TESTES	AAAA TTTT P	RRR ATIONAL TION THE HEILTH	ASS EX BER HO BER GI	Y Y Y H A	R R R W W L L W W W W W W W W W W W W W	H H H H H	THE COCCO COCCO	T X X X X X X X X X X X X X X X X X X X	D N H H P P	21 2 1 2 1 2 1 2 2 1 2 2 1 2 2 2 2 2 2	T T T T T T T T T T T T T T T T T T T	THY - LLLL H	FFF - S - L LLL S	HHH R - X RRXX E EE	AAP - T E EEOE L	S A L L L L L L L L L L L L L L L L L L	D I	A J A J A J A J A J A J A J A J A J A J	23 0 0 N - P 24 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	DDD S O S S N L O L L S R O R R	TTT R RRDH S SSSA I	DOC PPRO E E E	D D N S S T P R R R R P P	AAA SSSS H HHY L	K T S S S S S S S S S S S S S S S S S S		7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1		TO THE STATE OF TH	N N S T V	HHH P	YYYAAAASSDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	HHH O O ND P	YYY	REG - O OOK R	PPPL	KKK S 2 SSTN T TTTE	NNS BENN STORES OF STORE	TI T	heine	D.P. P.	RO R
148 150 146 199 198 200 196 238 237 250 230	RADRND RVN ELS ELS ELS DLS THN NHN TNT TOS	NAS AS AS LT LTSN RG	VVV E EEEE V VVV V VVV	A L A A A A A A A A A A A A A A A A A A	AAAA T TTTT P PPP	RRR ATIONALA Y TO YYME HILL HHH	A S S E J J J J J J J J J J J J J J J J J	THE HEAD IN THE HE	R R R W W L L W W W W W W W W W W W W W	H H H H H H H H H H H H H H H H H H H	THE COUCO O CONS	T X X X X X X X X X X X X X X X X X X X	D N N N N N N N N N N N N N N N N N N N	22 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3		TIV - L LLLL	777 S - L LLLL S - S 8 8	HIND R RREE	AAP - T- E BEOR L LLL	S S A L L L L L L L L L L L L L L L L L	D I I		23 0 0 N P 21 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	DDD SOSSH LOLLS RORRE	TTT R RRDH S SSSA I III	DOC P PRIQ E EMOO I III	D D N S S T P R R R R P P P P	AAA SSSS H HHHY L	K K S S S S S S S S S S S S S S S S S S	TANK TANK TANK TANK TANK TANK TANK TANK	7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1		TO STATE STA	N N S T V	H H H H H H H H H H H H H H H H H H H	YYAAAASDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	HHH O COND P	YYY	TIT R RRG - O OOOK R RRR	PPPL	K K S 2 S S T N T T T T T E B B E B	NNS BEENN STOLENS OF STORE		heinelinelinelinelinelinelinelinelineline	D.P. P.	RO R

6/17

Figure 6(b)

PSPOPTPEPSPSPOPA	PNAPSNPIDXKLV	KEAVRKVGDGYVFE	ENGV Hajority
360	370 380	3 å 0	400
337 PSPQSTPEPSPSPQPA	PNPQPAPSNPIDEKLV	KEAVRKVGDGYVPE	E-N G V PhtD. PRO
336 PSPQPTPEPSPSPQP		KEAVRKVG D G Y V F E	ENGV PhtB.pro
350 PSPOPTPEPSPGPOPA		O PIAK A CIEIO A A E E	
			_
SRYVPAKDLSAETAAG	LDSKLAKOESLSHKLG	AKKIDLPSSDREFY	N K A Y Hajority
410	420 430	440	450
387 S R Y I P A K D L S A E T A A G	IDSKLARQESLSHKLG	KKTDLPSSDREFY	
380 S RYTPAKOL SAETAAG 396 <u>S RY</u> VPAKOLPSETVKN	LIES & LAK Q E S L S H & L GIV	CK K T D L P S S D R E F Y	N K A Y PhtB.pro
318 G S TV S T N A K P N E V V S S	LGSL8	HPS SLT TS	
			F.L. B. Madamirus
DLLARIHODLLDHKGR	· · · · · · · · · · · · · · · · · · ·	•	•
460	470 480	490	500
437 DLLARIHQDLLDNKGR 430 DLLARIHODLLDNKGR	Q V D F E A L D R L L E R L E D . O V D F E A L D N L L E R L E D .	V S S D K V K L V D D I L A	FLAP Pheb.pro
446 NLLTEAHRALFENKOR	N SD POAL DELLER LND	B ST NKEK L V D D L L A	F L A P Phta. PRO
348		<u> </u>	Phtm.PRO
IRHPERLOXPHAGITY	TDDEIQVAKLAGKYTA	SDGYIFDPRDITSD	E G D A Majority
510	520 530	540	550
487 JIRKPERLGKPNAQITY	TODEIQVAKLAGKYTT	DOYIFDPRDITSD	EGDA PhtD.PRO
480 IRHPERLGEPHAQITY	TDDEIOVARLAGEYTA	EDGYIPD <u>PR</u> DI <u>T</u> SD	EGDA PhtB.pro
480 IRHPERLGKPHAQITY 496 ITHPERLGKPHSO 122	TODEIOVARLAGRYTÄ TEDEVRIAQLADKYTT	BOGYIPOPROITSO SOGYIPORROITSO	EGDA PhtB.pro EGDA PhtA.PRO
480 IRHPERLOKPHAGITY 496 ITHPERLOKPHSOIEY 353	TODETOVARLAGETTA TEDEVATAGLADKYTT	EDGYIFOPRDITSD SDGYIFDEHDINSD BDGYIFNPKDIVEE	EGDA PhtB.pro EGDA PhtA.PRO TATA PhtE.PRO
480 IRHPERLGKPHAQITY 496 ITHPERLGKPHSO 122	TODETOVARLAGETTA TEDEVATAGLADKYTT	EDGYIFOPRDITSD SDGYIFDEHDINSD BDGYIFNPKDIVEE	EGDA PhtB.pro EGDA PhtA.PRO TATA PhtE.PRO
480 IRHPERLOKPHAGITY 496 ITHPERLOKPHSOIEY 353	TODETOVARLAGETTA TEDEVATAGLADKYTT	EDGYIFOPRDITSD SDGYIFDEHDINSD BDGYIFNPKDIVEE	EGDA PhtB.pro EGDA PhtA.PRO TATA PhtE.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H M I K K D S	TODEIOVARLAGRYTA TEDEVRINGLADRYTT	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE GLTPPSTDKQDSGN S90	E G D A Phea.pro E G D A Phea.PRO T A T A Phee.PRO T E A K Majority 600 T E A K Pheb.PRO
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 517 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S	TODEIQVARLAGRYTA TEDEVRINQLADRYTT	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN	E G D A Phea.pro E G D A Phea.PRO T A T A Phee.PRO T E A K Hajority 600 T E A K Pheb.PRO T E A K Pheb.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H M I K K D S	T D D E I Q V A E L A G E Y T A TE D E V R I A Q L A D E Y T T A L S E A E R A A A O A Y A E E E S70	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE STORY S90 GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN	E G D A PheB.pro E G D A PheA.PRO T A T A PheE.PRO T E A K Hajority 600 T E A K PheD.PRO T E A K PheB.pro P T G D PheA.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353	T D D E I Q V A E L A G E Y T A TE D E V R I A Q L A D E Y T T A L S E A E R A A A O A Y A E E E S70	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE GLTPPSTDKQDSGN S90 GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN	E G D A Phea.pro E G D A Phea.PRO T A T A Phee.PRO T E A K Hajority 600 T E A K Pheb.PRO T E A K Pheb.PRO F T G D Phea.PRO S H E K Phee.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H M G H S H W I G K D S	T D D E I O V A E L A G E Y T A TE D E V R I A Q L A D K Y T T T L S E A E R A A A O A Y A K E K S70	EDGYIPDPRDITSD SDGYIPDEHDIISD EDGYIPNPRDIVEE SLTPPSTDKQDSGN 590 SLTPPSTDKQDSGN LTPPSTDKQDSGN SLTPPSTDKQDSGN LTPPSTDKQDSGN LTPPSTDKQDSGN LTPPSTDKQDSGN LTPPSTDKQDSGN	E G D A Phea.pro E G D A Phea.PRO T A T A Phee.PRO T E A K Hajority 600 T E A K Pheb.PRO T E A K Pheb.PRO F T G D Phea.PRO S H E K Phee.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H H T H S H W I K K D S 546 Y V T P H H G H S H W I G K D S 172 YI V R H G D H F H Y I P K G A E A I Y N R V K A A K K V P 610	T D D E I O V A E L A G E Y T A TED E V R I A Q L A D K Y T T L S E A E R A A A O A Y A K E K 570 580 L S E A E R A A A Q A Y A K E K L S E A E R A A A Q A Y A K E K L S E A E R A A A Q A Y A K E K L S D E E K V A A O A Y T K E R S H Q I G Q P T L P H H S L D R H P Y H L Q Y T V E V K H 620 630	EDGYIPDPRDITSD EDGYIPNPRDIVEE GLTPPSTDKQDSGN 590 GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN GLTPPSTDKQDSGN	E G D A PhtB.pro E G D A PhtA.PRO T A T A PhtE.PRO T E A K Hajority 600 T E A K PhtD.PRO T E A K PhtB.pro P T G D PhtA.PRO S H E K PhtB.PRO K F E W Hajority 650
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H H G H S H W I G K D S 372 Y I V R H G D H F H Y I P K G A E A I Y N R V K A A K K V P 587 G A E A I Y N R V K A A K K V P	T D D E I O V A E L A G E Y T A TE D E V R I A Q L A D K Y T T L S E A E R A A A O A Y A K E K S70	EDGYIPDPRDITSD BDGYIPDBHDIISD BDGYIPNPRDIVEE GLTPPSTDHQDSGN S90 GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSTDHQDSGN GLTPPSPDADVKAN LATPSPSLPINPGT GSLIIPHYDHYHNI	E G D A Phea.pro E G D A Phea.pro T A T A Phea.pro T E A K Hajority 600 T E A K Pheb.pro T E A K Pheb.pro P T G D Phea.pro S H E K Phee.pro K F E W Hajority 650 K F E W Pheb.pro
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S O I E Y 353	T D D E I Q V A E L A G E Y T A TE D E V R I A Q L A D E Y T T A L S E A E R A A A O A Y A E E E 570	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE SLTPPSTDHQDSGN CLTPPSTDHQDSGN CLTPPSPDADVXAM CATPSPSLPINPGT GSLIIPHYDHYHMI GSLIIPHYDHYHMI	E G D A PheB.pro E G D A PheA.PRO T A T A PheE.PRO T E A K Hajority 600 T E A K PheD.PRO T E A K PheB.pro P T G D PheA.PRO S H E K PheB.PRO K F E W Hajority 650 K F E W PheB.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H H G H S H W I G K D S 372 Y I V R H G D H F H Y I P K G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P	T D D E I Q V A E L A G E Y T A TE D E V R I A Q L A D E Y T T A L S E A E R A A A O A Y A E E E 570	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE SLTPPSTDHQDSGN CLTPPSTDHQDSGN CLTPPSPDADVXAM CATPSPSLPINPGT GSLIIPHYDHYHMI GSLIIPHYDHYHMI	E G D A PheB.pro E G D A PheA.PRO T A T A PheE.PRO T E A K Hajority 600 T E A K PheD.PRO T E A K PheB.pro P T G D PheA.PRO S H E K PheB.PRO K F E W Hajority 650 K F E W PheB.PRO
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S O I E Y 353	T D D E I Q V A K L A G K Y T A TE D E V R I A Q L A D K Y T T	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE SPO GLTPPSTDHQDSGN GLTPPSPDADVKAN GLTPSPSLPINPGT GSLIIPHYDHYHNI GSLIIPHYDHYHNI GNLIIPHKDHYHNI GNLIIPHKDHYHNI GNLIIPHKDHYHNI	E G D A Phea.pro E G D A Phea.PRO T A T A Phee.PRO T E A K Hajority 600 T E A K Pheb.PRO T E A K Pheb.PRO F T G D Phea.PRO S H E K Phea.PRO K F E W Hajority 650 K F E W Pheb.PRO K F E W Pheb.PRO K F A W Pheb.PRO K F A W Pheb.PRO F A W Pheb.PRO F A W Pheb.PRO F A W Pheb.PRO
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S O I E Y 353	T D D E I Q V A K L A G K Y T A TE D E V R I A Q L A D K Y T T	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE SPO GLTPPSTDHQDSGN GLTPPSPDADVKAN GLTPSPSLPINPGT GSLIIPHYDHYHNI GSLIIPHYDHYHNI GNLIIPHKDHYHNI GNLIIPHKDHYHNI GNLIIPHKDHYHNI	E G D A Phea.pro E G D A Phea.PRO T A T A Phee.PRO T E A K Hajority 600 T E A K Pheb.PRO T E A K Pheb.PRO F T G D Phea.PRO S H E K Phea.PRO K F E W Hajority 650 K F E W Pheb.PRO K F E W Pheb.PRO K F A W Pheb.PRO K F A W Pheb.PRO F A W Pheb.PRO F A W Pheb.PRO F A W Pheb.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H H G H S H W I G K D S 546 Y V T P H H G H S H W I G K D S 572 Y I V R H G D H F H Y I P R G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 596 S A A A I Y N R V K A A K K V P 596 S A A A I Y N R V K A A K K V P 417 H E E D G Y G P D E G L Y E A P K G Y T L E D 660	TODEIOVARLAGRYTA TEDEVRIAQLADKYTT LSEAERAAAOAYAKEK \$70	EDGYIPDPRDITSD EDGYIPDEHDIISD EDGYIPNPRDIVEE STORY IPNPRDIVEE STORY IPNPRDIVEE STORY IPNPRDIVEE STORY IPNPRDIVEE STORY INNPRDIVEE STORY INNER STORY IN	E G D A Pheb.pro E G D A Pheb.pro E G D A Pheb.pro T E A K Hajority 600 T E A K Pheb.pro T E A K Pheb.pro F T G D Pheb.pro S H E K Pheb.pro K F E W Hajority 650 K F E W Pheb.pro K F A W Pheb.pro
480 I R H P E R L G X P N A Q I T Y 496 ITH P E R L G X P N S O I E Y 353 Y V T P H H T H S H W I K K D S 560 517 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H H G H S H W I G K D S 772 Y I V R H G D H F H Y I I P K G A E A I Y N R V K A A K K V P 587 G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 596 S A A A I Y N R V K A A K K V P 596 S A A A I Y N R V K A A K K V P 417 H E E D G Y G F D E G L Y E A P K G Y T L E D 660 637 F D E G L Y E A P K G Y T L E D 630 P D E G L Y E A P K G Y T L E D	T D D E I Q V A K L A G K Y T A TE D E V R I A Q L A D K Y T T	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE BDGYIPNPRDIVEE BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSPDADVKAN BATPSPSLPINPGT BSLIIPHYDHYHMI BGSLIIPHYDHYHMI BGFVMSHGDHNH HSDNGFGNASDHVK 690	E G D A PheB.pro E G D A PheB.pro E G D A PheB.pro F A T A PheB.pro T E A K Hajority 600 T E A K PheB.pro F T G D PheB.pro F T G D PheB.pro K F E W Hajority 650 K F E W PheB.pro K F E W PheB.pro K F A W PheB.pro
480 I R H P E R L G X P N A Q I T Y 496 ITH P E R L G X P N S O I E Y 353 Y V T P H H T H S H W I X X D S 560 537 Y V T P H H T H S H W I X X D S 546 Y V T P H H T H S H W I X X D S 546 Y V T P H H G H S H W I G X D S 372 Y I V R H G D H F H Y I P R G A E A I Y N R V K A A K X V P 580 G A E A I Y N R V K A A K X V P 580 G A E A I Y N R V K A A K X V P 596 S A A A I Y N R V K A A K X V P 596 S A A A I Y N R V K A A K X V P 417 H E E D G Y G F D E G L Y E A P K G Y T L E D 660 637 F D E G L Y E A P K G Y T L E D	T D D E I Q V A K L A G K Y T A TE D E V R I A Q L A D K Y T T	EDGYIPDPRDITSD BDGYIPNPRDIVEE BDGYIPNPRDIVEE BDGYIPNPRDIVEE BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSTDHQDSGN BLTPPSPDADVKAN BATPSPSLPINPGT BSLIIPHYDHYHMI BGSLIIPHYDHYHMI BGFVMSHGDHNH HSDNGFGNASDHVK 690	E G D A PheB.pro E G D A PheB.pro E G D A PheB.pro F A T A PheB.pro T E A K Hajority 600 T E A K PheB.pro F T G D PheB.pro F T G D PheB.pro K F E W Hajority 650 K F E W PheB.pro K F E W PheB.pro K F A W PheB.pro

7/17

Figure 6(c)

XXXXXXXXX	XXXEEX -	PE		• • • • • • • •	Hajor
	710	720 .	730	740	750
87 DQDSKPDED	KEHDEVSEP	THPESDEKE	NHAGLNPSADN		
SO COADTHOTE		EKPEET	PREEK	PQSEKPESPI	
96 H S EDPWK N F	KY DE EL				· PhtA. · PhtE.
149					
EEXEETPXE	XEXPOVETE	KVEAKLXEAI	EXLLXXVTDPS	IKXNAXETLI	G L Hajor
	760	770	780	790	800
37 E E A E DIT DE	AEIPQVENS	VINAKIADA	EALLEKVTDPS	IR ON A HET L	G L PhtD.
22 EPBESPEE	SEEPQVETE	XVEEKLREA	FOR LOKEODPE	IKSNAKETLI	G L Phes.
	PEVPOVETE	- PKKDLTEZ	EVLLAKVTDSS	LKANATETLI	l <u>G L</u> PhtA. PhtE.
149		- F K P M T ELE (PACE.
KNNLLLGTX	DNHTIHAEA	EKLLALLKES	<u> </u>	- K	Hajor
;	sio	820	830	840	
87 KSSLLLGTK					PhtD.
72 KNN L L F G T Q		EKLLALLKE			PhtB.
	ID N NISII N Y R Y	EKLLALLKIGU	SHPSSVSKEKI	RENI	PhtA.
59				·· [0] ·· ·	*****

Decoration 'Decoration #2': Box residues that match the Consensus exactly

Figure 7(a)

	TCCTATG	AGCTTGGAL	JTATCAAGC	TGGTCAGGTT	AAGAAAGAGT	C T A A Hajority
		10	20	30	40	50
	TCTTACE	AGTTGGGAC	TGTATCAAGC	TAGAACGGTT	AAGGAAAA	T A A phtA.SEO
61 1	TCCTATG	AGCTTGGAC	GTTACCAAGC	TGGTCAGGAT	AAGAAAGAGT	CTAA phtB.seq
i	TCCTATG.	AACTTGGTC	GTCACCAAGC	TGGTCAGGTT	AAGAAAGAGT	CTAA phtD.SEQ
64	GCCTATG	C	AGCATC GT	TCG-CAGGAA	A A T A A G G A C A	ATAA phte.SEQ
	TCGTGTT	<u> </u>	GATGGTGATC	AGGCTGGTCA	******	A A C T HAJORITY
		60	70	80	90	100
108				AAGCGACGCA		
51	TCGAGTT	GCTTATATA	GATGGTGATG	AGGCTGGTCA	A A A G G C A G A A	AACT phtB.seq
51				AGGCTGGTCA		
111	TCGTGTC	TCTTATGT6	GATGGCAGCG	AGTCAAGTCA		A A C I phillips
	TGACACC	AGATGAGGT	TAGTAAGAG	GAGGGATCA	ACGCTGAGCA	A A T T Majority
		110	120	130	140	150
		- 1	T10C116C01	GAAGGAATCA	ATGCTGAGCA	AATC phta.SEO
158 101	TGACTCC	l G A T G A A G T	CAGTAAGAG	GAGGGGATCA	ACGCCGAACA	AATT phtB.seq
101	TGACACC	AGATGAAGT	CAGTAAGAGG	; G	. * C C C C C * * * C *	AATC phtD.SEQ
161	TGACACC	AGACCAGGT	TAGCCAGAA		AGGCTGAGCA	AATT pht2.SEQ
				TGTGACCTCT	C	A T T A Majority
	GTCATCA	AGATTACCO	X T C X X O O I I I		•	•
		160	170	180	190	200
208	GTCATCA	AGATAACAG	ACCAAGGCT	TGTCACTTCA	CATGGCGACC	A C T A phtA.SEQ
151	GTTATCA	AGATTACGG	ATCAAGGTT	LTGTGACCTCT	'C	ATTA phtB.seq
151	GTTATCA		ATCAAGGTT!	LTGTGACCTCT LTGTGACCTCT	CATGGAGACC CATGGAGACC	ATTA phtB.seq ATTA phtD.SEQ
151	G T T A T C A G T C A T C A G T A A T C A	A G A T T A C G G A G A T T A C G G A A A T T A C A G	ATCAAGGTT; ATCAAGGTT; ATCAGGGCT;	\	C	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ
151	G T T A T C A G T C A T C A G T A A T C A	A G A T T A C G G A G A T T A C G G A A A T T A C A G	ATCAAGGTT; ATCAAGGTT; ATCAGGGCT;	LTGTGACCTCT LTGTGACCTCT	C	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ
151	G T T A T C A G T C A T C A G T A A T C A	A G A T T A C G G A G A T T A C G G A A A T T A C A G	ATCAAGGTT; ATCAAGGTT; ATCAGGGCT;	\	C	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ
151 151 211	GTTATCA GTCATCA GTAATCA TCATTAC	A G A T T A C G G A G A T T A C G G A A A T T A C A G T A T A A T G G G T A C A A T G G T	ATCAAGGTTAATCAAGGTTAATCAAGGTTCCT	T G T G A C C T C T A T G T G A C C T C T A T G T A A C G T C A E A T G A T G C C A T 230 F A T G A C G C T A T	CATGGAGACC CATGGAGACC CACGGTGACC CATCAGTGAA 240	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ
151 151 211	GTTATCA GTCATCA GTAATCA TCATTAC TCATTAT	A G A T T A C G G A G A T T A C G G A A A T T A C A G T A T A A T G G G T A C A A T G G G	ATCAAGGTTAATCAAGGTTCATCAAGGTTCCT	TGTGACCTCT TGTGACCTCT TGTAACGTCA CATGATGCCAT 230 CATGACGCTAT CATGACGCTAT	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq
151 151 211	GTTATCA GTCATCA GTAATCA TCATTAC TCATTAT TCATTAC	A G A T T A C G G A G A T T A C G G A A A T T A C A G T A T A A T G G G T A C A A T G G G T A T A A T G G G	ATCAAGGTTAATCAAGGTTCATCAAGGTTCCT	TGTGACCTCT TGTGACCTCT TGTAACGTCA CATGATGCCAT 230 FATGACGCTAT FATGACGCTAT FATGATGCCAT	CATGGAGACC CATGGAGACC CACGGTGAC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq
151 151 211 258 201	GTTATCA GTCATCA GTAATCA TCATTAC TCATTAT TCATTAC	A G A T T A C G G A G A T T A C G G A A A T T A C A G T A T A A T G G G T A C A A T G G G T A T A A T G G G	ATCAAGGTTAATCAAGGTTCATCAAGGTTCCT	TGTGACCTCT TGTGACCTCT TGTAACGTCA CATGATGCCAT 230 CATGACGCTAT CATGACGCTAT	CATGGAGACC CATGGAGACC CACGGTGAC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq
151 151 211 258 201 201	GTTATCA GTCATCA GTAATCA TCATTAC TCATTAT TCATTAC TCATTAC TCATTAC	AGATTACGGAAATTACGGAAATTACAGC 210 TACAATGGTATACAGTACAGTATAATGGTATAATGGCTATAATGGCTATAATGGCTATAATGGCTATAATGGC	ATCAAGGTTAATCAAGGTTCATCAAGGTTCCT	TGTGACCTCT TGTGACCTCT TGTAACGTCA 230 TATGACGCTAT TATGACGCTAT TATGACGCTAT TATGATGCCAT	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA CATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtB.seq
151 151 211 258 201 201	GTTATCA GTCATCA GTAATCA TCATTAC TCATTAT TCATTAC TCATTAC TCATTAC	A G A T T A C G G A G A T T A C G G A A A T T A C A G C C C C C C C C C C C C C C C C	ATCAAGGTTAATCA	TGTGACCTCT ATGTGACCTCT ATGTAACGTCA 230 FATGACGCTAT FATGACGCTAT FATGACGCTAT FATGACGCTAT FATGATGCCAT FATGATGCCAT	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtC.SEQ GAAC phtE.SEQ
151 151 211 258 201 201 261	TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC	A G A T T A C G G A G A T T A C G G A A A T T A C A G C C T A T A A T G G G T A T A A T G G G T A T A	ATCAAGGTTAATCA	TGTGACCTCT TGTGACCTCT ATGTAACGTCA 230 FATGACGCTAT FATGACGCTAT FATGACGCTAT FATGACGCTAT FATGATGCCAT FATGATGCCAT FATGATGCCAT FATGATGCCAT	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA CATCAGTGAA CCATCAGTGAA CCATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtD.SEQ GAAC phtE.SEQ CAAT Majority 300
151 151 211 258 201 201 261	TACTCAT GTTATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC	A G A T T A C G G A G A T T A C G G A A A T T A C A G C C T A T A A T G G G T A T A A T G G G T A T A	ATCAAGGTTAATCA	TGTGACCTCT TGTGACCTCT TGTAACGTCA 230 PATGACGCTAT TATGACGCTAT TATGATGCCAT TATGATGCCAT TATGATGCCAT TATGATGCCAT TATGATGCCAT TATGATGCCAT	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CTTTAGTGAA CTTTAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtB.seq CAAT Majority 300 TAAT phtA.SEQ
151 151 211 258 201 201 261	TACTCAT	A G A T T A C G G A G A T T A C G G A A A T T A C A G C T A T A A T G G G T A T A A T G G G T A T A	ATCAAGGTTAACAAGGTTAACAAGGTTCCT	TGTGACCTCT ATGTGACCTCT ATGTAACGTCA 230 FATGATGCCAT ATGATGCCAT TATGATGCCAT	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtB.seq CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq
151 151 211 258 201 201 261	TACTCAT TCCTCAT TCCTCAT	A G A T T A C G G A G A T T A C G G A A A T T A C G G C T A T A A T G G G T A T A A T G G G T A T A	ATCAAGGTTAAGGTTAAGGTTCAAGGTTCCTAAGGTTCCTAAAGGTTCCTAAAGGTTCCTAAAGGTTCCTAAAGGTTCCTAAAGGTTCCTAAAGGTTCCTAAAGGAATTATCAAGAATTATCAAGAATTATCAAGAATTATCA	TGTGACCTCT ATGTGACCTCT ATGTAACGTCA 230 FATGATGCCAT TATGATGCCAT TATGATGCAT TATGAT TATGATGCAT TATGAT TATGATGCAT TATGATGCAT TATGATGCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT T	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtB.seq CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.SEQ CAAT phtB.SEQ
151 151 211 258 201 201 261	TACTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT	AGATTACGG AGATTACGG AAATTACAG 210 TACAATGGT TATAATGGC TATAATGGC TATAATGGC TATAATGGC TATAATGGC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC	ATCAAGGTTAATCAGATTATCAGAATTATC	TGTGACCTCT ATGTGACCTCT ATGTGACCCAT 230 FATGATGCCAT GTTGAAGGATT GTTGAAGGATT GTTGAAGGATT ACTTAAAGACG	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CCAGACATTGT CAGACATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtD.SEQ GAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq
151 151 211 258 201 201 261	TACTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT	AGATTACGG AGATTACGG AAATTACAG 210 TACAATGGT TATAATGGC TATAATGGC TATAATGGC TATAATGGC TATAATGGC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC	ATCAAGGTTAATCAGATTATCAGAATTATC	TGTGACCTCT ATGTGACCTCT ATGTAACGTCA 230 FATGATGCCAT TATGATGCCAT TATGATGCAT TATGAT TATGATGCAT TATGAT TATGATGCAT TATGATGCAT TATGATGCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT TATGATCAT T	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CCAGACATTGT CAGACATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtD.SEQ GAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq
151 151 211 258 201 201 261	TACTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT	AGATTACGG AGATTACGG AAATTACAG 210 TACAATGGT TATAATGGC TATAATGGC TATAATGGC TATAATGGC TATAATGGC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC	ATCAAGGTTAATCAGATTATCAGAATTATC	TGTGACCTCT ATGTGACCTCT ATGTGACCCAT 230 FATGATGCCAT GTTGAAGGATT GTTGAAGGATT GTTGAAGGATT ACTTAAAGACG	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CCAGACATTGT CAGACATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtD.SEQ GAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq
151 151 211 258 201 201 261	GTTATCA GTCATCA GTCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT TCCTCAT	AGATTACGG AGATTACGG AAATTACAG 210 TACAATGGT TATAATGGC TATAATGGC TATAATGGC TATAATGGC TATAATGGC TATAATGGC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC GAAAGATCC	ATCAAGGTTA ATCAAGGTTA ATCAAGGTTCCT 220 AAGGTTCCT AAGGTTCCT AAGGTTCCT AAGGTTCCT AAGGTTCCT AAAGTTCCT AAAGTTCCT AAAGTTCCT AAAGTTCCT AAAGTTCCT AAAGTTCCT AAAGTTCCA 320 ATGTTATCA	TGTGACCTCT ATGTGACCTCT ATGTAACGTCA 230 PATGACGCTAT ATGATGCCAT TATGATGCCAT TATGATGCAT TATGATGCAT TATGATGATGGAT TATGATGATGGAT TATGATGATGGAT TATGATGGATG	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CCATCAGTGAA CCAGACATTGT CCAGACATTGT CCAGACATTGT CCAGACATTGT CCAGACATTGT CAGACATTGT CAGACATTGT CAGACATTGT CAGACATTGT CAGACATTGT CAGACATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtB.seq CAAT Majority 100 TTAAT phtA.SEQ CAAT phtB.seq
151 151 211 258 201 201 261 308 251 251 311	GTTATCA GTCATCA GTAATCA TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC GAGGTCA GAGGTCA GAGATCA	A G A T T A C G G A G A T T A C G G A A A T T A C G G T A T A A T G G G T A T A A T G G G A A A G A T C G G A A A G A T C G G A A A G A T C G G A A G A T C G G A A G G T G G T T T T T T T T T T T	ATCAAGGTTA ATCAAGGTTA ATCAAGGTTCT 220 AAGGTTCCT AAGGTTCCT AAGGTTCCT AAGGTTCCT AAGGTTCCT AAGGTTCCT AAAGTTCCT AAAGTTCCT AAAGTTATCA GAATTATCA AAACTATAA AAACTATAA AAACTATAA AAACTATAA AAACTATAA AAACTATAA AAACTATAA AAACTATAA	TGTGACCTCT ATGTGACCTCT ATGTAACGTCA 230 PATGACGCTAT ATGATGCCAT TATGATGCCAT TATGATGCAT TATGATGCAT TATGATGATGCAT TATGATGCAT TAT	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CCATCAGTGAA CCAGACATTGT CCAGACATTGT CCAGACATTGT CCAGACATTGT CCAGACATTGT CAGACATTGT CAGACATTGT CAGACATTGT CAAAATACTATCT AAAATACTATCT CAAAATACTATCT CAAAATACTATCT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtB.seq CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq TTTTA Majority 350 TTTTA Majority
151 151 211 258 201 201 261 308 251 251 311	GTTATCA GTCATCA GTAATCA TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC TCATTAC GAGGTCA GAGGTCA GAGATCA GAAATCA	AGATTACGGAAATTACGGAAATTACAGGTATAATGGCTATAATGGCTATAATGGCAAAAGATCCGAAAAGATCCGAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAGATCCGAAAAAAAA	ATCAAGGTTA ATCAAGGTTA ATCAAGGTTCTT 220 AAGGTTCCTT AAGGTTCCTT AAGGTTCCTT AAGGTTCCTT AAGGTTCCTT AAGGTTCCTT AAAGTTCCTT AAAGTTCCTT AAAGTTCCTT AAAGTTATCA GAATTATCA GAATTATCA ATGTTATCA ATGTTATCA	TGTGACCTCT ATGTGACCTCT ATGTAACGTCA 230 PATGACGCTAT ATGATGCCAT TATGATGCCAT TATGATGCAT TATGATGCAT TATGATGATGGAT TATGATGATGGAT TATGATGATGGAT TATGATGGATG	CATGGAGACC CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCATCAGTGAA CCAGACATTGT CCAGACATTGT CCAGACATTGT CCAGACATTGT CCAGACATTGT CAGACATTGT CAGACATTATTGT CAGACATTATTATTATTATTATTATTATTATTATTATTATTAT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtB.seq GAGC phtB.seq CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq TTTA Hajority 150 TTTTA Hajority 150 TTTTA phtA.SEQ TTTTA phtB.seq

9/17

Figure 7(b)

160 310 310 310 320 400		CCTTAAGGATGCAGL	CATGCGGATAA	TGTTCGGAC	***	A T T A Majority
		360	370	380	390	400
	408	CCTTAAGGATGCTGC	CACGEGGATAA	CGTCCGTAC	AAAGAGGAA	A T C A phta.SEO
### ##################################						•
### ATCOTCAGRAGCAGGAACATAGTCATAATCATGAGGGTGGAXCTA Majority ####################################	351					
410 420 430 440 450 458 ATCGACAAAAACAAGAGCATAGTCAACATCGTGAAGATGGAACTCCAAGAA phth.seq 401 AACGTCAGAAAACAAGAGCATAGTCATAACATCGTGAAGATGGAACTCCAAGAA phth.seq 401 AACGTCAGAAAACAAGAACACAGTCATAATCATAACTCAAGAGCA phth.seq 401 AACGTCAGAAACCAAGAACACAGTCATAATCATAACTCAAGAGCA phth.seq 401 AACGTCAGAAACCAAGAACACAGTCATAATCATAACTCAAGAGCA	411	CCTGAAAGATGCAGC	CATGCTGATAA	LTGTTCGAAC	TAAAGATGAA	ATCA phtE.SEQ
410 420 430 440 450 458 ATCGACAAAAACAAGAGCATAGTCAACATCGTGAAGATGGAACTCCAAGAA phth.seq 401 AACGTCAGAAAACAAGAGCATAGTCATAACATCGTGAAGATGGAACTCCAAGAA phth.seq 401 AACGTCAGAAAACAAGAACACAGTCATAATCATAACTCAAGAGCA phth.seq 401 AACGTCAGAAACCAAGAACACAGTCATAATCATAACTCAAGAGCA phth.seq 401 AACGTCAGAAACCAAGAACACAGTCATAATCATAACTCAAGAGCA		ATCGTCAGAAGCAGG.		ATCATGAGG	GTGGAXCT	- A Majority
ATEGRACARARACAR GROCATAGTCARCATEGTGRAGGTGGRACTCCARGA phth.seq						
A			<u></u>			
A A C G T C A A A A C C A G T C A T C A T A T C A C G G G G G T G T C T						
ATCGTCAAAAACAAGAACATGTCAAAAGATAATGAG						
						G G G L Wadandan
A		GATGATXXTGCTGTT	JETOT KOCC KOK	CTECEXXBOX	CGCTATACAA	C G G X RAJORITY
445 AACGATCAAGCAGTAGCTAGCAGCCAAAGCCCAAAGACGTTATACAACGGA phtb.seq 445 AACGATCAAGCAGTAGTTGCAGCCAAGGCCAAAGGACGCTTATACAACGGA phtb.seq 499 GTTAACCTCTAATGTTGCTGTAGCAAGGTCTAAGGACGCTATACAACGACAAA phtb.seq 499 GTTAACTCTAATGTTGCTGTAGCAAGGTCTCAAGGACGCTATACAACAAA phtb.seq 490 GTTAACTCTAATGTTGCTGTAGCAAGGTCTCAAGGACGATATACAACAAA phtb.seq 491 TGATGGTTATATCTTTAATGCATCTGATATCATTGAGGATACGGGTGATG Majority 510 520 530 540 550 492 TGATGGGTATATCTTTAATGCTTCTGATATCATTGAGGACACGGGTGATG phtb.seq 495 TGATGGGTATATCTTCAATGCATCTGATATCATTGAGGACACGGGTGATG phtb.seq 495 TGATGGGTATATCTTCAATGCATCTGATATCATTGAGGACACGGGTGATG phtb.seq 496 TGATGGTTATATCTTCAATGCATCTGATATCATTGAGGACACGGGTGATG phtb.seq 547 CTTATATCGTTCCTCAATGCGATCATTACCATTACATTCCTAAGGAATGAG phtb.seq 548 CTTATATCGTTCCTCAATGGAGATCATTACCATTACATTCCTAAGGAATGAG phtb.seq 549 CTTATATCGTTCCTCAACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 540 570 580 590 600 541 CTTATATCGTTCCTCAACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 542 CTTATATCGTTCCTCAACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 543 CTTATATCGTTCCTCAACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 544 CTTATATCGTTCCTCAACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 545 CTTATATCGTTCCTCAACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 546 610 630 640 650 557 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtb.seq 558 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtb.seq 559 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtb.seq 550 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGA phtb.seq 551 TTATCAGCTAGCGAGTTAGCTGCTGCTGCAGAAAGCCATCTGGCAGAA phtb.seq 552 TTATCAGCTAGCGAGTTAGCTGCTGCTGCAGAAAGCCATCTGGCAATAGACCTT 564 TATCAGCTAGCGAGTTAGCTGCTGCTGCTGCTTCCTAAGTTATAACCTT 565 TTATCAGCTAGCGAGTTAGCTGCTGCTTCCTTCAAGTTCTAAGTTATAACTT 566 670 680 690 700 667 AAATCTGTCAAATTCAAGAACCTATCGCCGCGACAAAAATAGCGATTAAAATACCTT 570 AAATCTGTCAAATTCAAGAACCTTTCTTCAAGTTATAAATG C phtb.seq 668 AAATCTGTCAAATTCAAGAACCTTTCTTCTTCAAGTTATAAATG C phtb.seq 668 AAATCTGTCAAATTCAAGAACCTTTCTTCTTCTAAGTTATAAATG C phtb.seq 668 AAATCTGTCAAATTCAAGAACCTTTCTTCTTCTAAGTTATAAATG C phtb.s		460	470	480	490	500
A A C G A T C A A G C A G T A G C A G C C A A G G A C G C T A T A C A A C G A Phtc. SEQ	508					
TGATGGTTATTCTTTAATGCTGTAGCAAGGTCTCAGGGACGATATACGACAAA phte.SEQ TGATGGTTATATCTTTAATGCATCATGATATCATTGAGGATACGGGTGATG Majority S10						
TGATGGTTATATCTTTAATGCATCTGATATCATTGAGGATACGGGTGATG Kajority						
Side	477	o : : xxc : c : xx : c : .				
TO AT GOTTATAT CTTTAAT GCTT CTGATAT CATAGAGGAT ACTGGTGATG PhtA.SEQ 495 TGATGGTATAT CTTCAATGCAT CTGATAT CATAGAGGAT ACTGGTGATG PhtB.seq 495 TGATGGTATAT CTTCAATGCAT CTGATAT CATTGAGGACACGGGTGATG PhtB.seq 495 TGATGGTTATAT CTTCAATGCAT CTGATAT CATTGAGGACACGGGTAATG PhtB.SEQ 496 TGATGGTTATAT CTTCAATGCAGCTGATAT TACATTGCAAGAATGAG PhtB.SEQ 497 S80 S90 600 498 CTTATAT CGTTCCTCATGGAGAT CATTACCATTACATTCCTAAGAATGAG PhtB.SEQ 498 CTTATAT CGTTCCTCATGGAGAT CATTACCATTACATTCCTAAGAATGAG PhtB.SEQ 499 CTTATAT CGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG PhtB.SEQ 490 S90 CTTATAT CGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG PhtB.SEQ 490 CTTATAT CGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG PhtB.SEQ 490 CTTATAT CGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG PhtB.SEQ 491 CTTATATCGTTCCTCATGGAGGTCACTATCACTTCCTAAGAATGAG PhtB.SEQ 492 CTTATATCGTTCCTCATGGAGGTCACTATCACTTACATTCCTAAGAATGAG PhtB.SEQ 493 CTTATATCGTTCCTCATGGAGGTCACTATCACTTACATTCCCAAAAGCGAT PhtB.SEQ 494 CTTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA PhtB.SEQ 495 CTTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGA PhtB.SEQ 495 CTTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGA PhtB.SEQ 496 CTTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGA PhtB.SEQ 497 CTTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGA PhtB.SEQ 406 CTTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGA PhtB.SEQ 407 CTTATCAGCTAGCGAGTTAGCTGCTGCTGCAGAAAGCCTATTGGAATGGAA		TGATGGTTATATCTT	TAATGCATCTG	LTATCATTOA	GGATACGGGT	G A T G Majority
### ##################################		510	520	E30	540	550
TGATGGGTATATCTTCAATGCATCTGATATCATTGAGGACACGGGTGATG phtb.seq 495 TGATGGTTATATCTTCAATGCATCTGATATCATTGAGGACACGGGTGATG phtb.seq 549 TGATGGTTATATCTTCAATGCATCTGATATCATTGAGGACACGGGTAATG phtb.seq 640 TGATGGTTATGCTTTAATCCAGCTGATATTATCGAAGATACGGGTAATG phtb.seq 650 \$70 \$50 \$90 600 650 CTTATATCGTTCCTCATGGAGATCATTACCATTACATTCCTAAGAATGAG phtb.seq 641 AGATGGTTCCTCATGGAGATCATTACCATTACATTCCTAAGAATGAG phtb.seq 652 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 653 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 654 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.seq 655 CTTATATCGTTCCTCATGGAGGTCACTATCACTACATTCCTAAGAATGAG phtb.seq 656 CTTATATCGTTCCTCATGGAGGTCACTATCACTACATTCCTAAGAATGAG phtb.seq 657 CTTATATCGTTCCTCATGGAGGTCACTACCATACATTCCCAAAAGCGAT phtb.seq 658 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtb.seq 659 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTCCTATTTGGAATGGGA phtb.seq 650 650 650 650 700 660 670 680 690 700 661 AGATCTGTCAAATTCAAGAACCTATCCGCGACAAAATAGCGATAAACACTT phtb.seq 663 AGATCTGTCAAATTCAAGAACCTATCCGCCGACAAAATAGCGATAAAACACTT phtb.seq 664 AGATCTGTCAAATTCAAGAACCTATCCGTCCTTCTTCAAGTTCTAGTTATAAATG phtb.seq 664 AGATCTGTCAAATTCAAGGAACCTATCCGTCCTTCTTCAAGTTCTAGTTATAAATG phtb.seq 664 AGATCTGTCAAATTCAAGGAACCTATCCGTCCTTCTTCAAGTTTCTAGGTTATAAATG phtb.seq 664 AGATCTGTCAAATTCAAGGAACCTATCCGTCCTTCTTCTAAGTTTATAAATG phtb.seq 665 AGATCTCTCTTCTTCTTCTAAGGTTCTAAGTTCTAAGTTATAAATG phtb.seq				330		
CTTATATCGTTCCTCATGGCGATCATTACCATTACCATTACCTAAGAATGAG Najority S60 S70 S80 S90 600 CTTATATCGTTCCTCATGGCGATCATTACCATTACCATTCCTAAGAATGAG Najority S60 S70 S80 S90 600 CTTATATCGTTCCTCATGGAGATCATTACCATTACATTCCTAAGAATGAG phtb.SEO CTTATATCGTTCCTCATGGAGATCATTACCATTACATTCCTAAGAATGAG phtb.SEO S70 S80 S90 600 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtb.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTTGGATGGGA Hajority 610 620 630 640 650 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTTGGATGGGA Phtb.SEO S70 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA Phtb.SEO S71 TATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA Phtb.SEO S72 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA Phtb.SEO S73 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA Phtb.SEO S74 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA Phtb.SEO S75 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA Phtb.SEO S75 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCATCTGGCTGGAA Phtb.SEO S76 TTATCAGCTAGCGAGTTAGCTGCTGCTTCTTCAAGTTCTAGTTATACTT Hajority S60 670 680 690 700 SAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT Phtb.SEO S70 AAATCTGTCAAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT Phtb.SEO S61 AGCAGGGATCTCGTCCTTCCTTCAAGTTCTAGTTATAATG Phtb.SEO	558	TGATGGTTATATCTT	TAATGCTTCTG			GATG phtA.SEO
CTTATATCGTTCCTCATGGCGATCATTACCATTACATTCCTAAGAATGAG Majority 560 570 580 590 600 608 CTTATATCGTTCCTCATGGAGATCATTACCATTACATTCCTAAGAATGAG phta.SEO 542 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phta.SEO 545 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phta.SEO 546 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phta.SEO 547 CTTATATCGTTCCTCATGGAGGTCACTATCACTACATTCCCAAAAGCGAT phte.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTTGGATGGGA Majority 610 620 630 640 650 658 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phta.SEO 592 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTATTGGAATGGGA phtb.SEO 659 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtb.SEO 649 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtb.SEO 640 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGCCAAAATAGCGAATACACTT phta.SEO 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtb.seo 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtb.seo		TGATGGGTATATCTT		ATATCATAGA ATATCATTGA	G G A T A C T G G T G G A C A C G G G T	GATG phtB.seq
S60 S70 S80 S90 600	492 495	TGATGGGTATATCTT	C	. T	G G A T A C T G G T G G A C A C G G G T G G A C A C G G G T	G A T G phtB.seq G A T G phtD.SEQ
CTTATATEGTTCCTCATGGAGATCATTACCATTACATTCCTAAGAATGAG phth.SEO 542 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phth.seo 545 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phth.seo 546 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phth.seo 547 CTTATATCGTTCCTCATGGAGGTCACTATCACTACATTCCTAAGAATGAG phth.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTTGGATGGGA Hajority 610 620 630 640 650 658 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phth.SEO 592 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phth.SEO 659 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phth.SEO 649 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCATCTGGCATGGGA phth.SEO 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA phth.SEO 660 670 680 690 700 708 AAATCTGTCAAAATTCAAGAACCTATCGCCGACAAAAATAGCGATAACACTT phth.SEO 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phth.seo 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phth.seo	492 495	TGATGGGTATATCTT	C	. T	G G A T A C T G G T G G A C A C G G G T G G A C A C G G G T	G A T G phtB.seq G A T G phtD.SEQ
CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtB.seq 545 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtD.SEO 599 CTTATATCGTTCCTCATGGAGGTCACTATCACTACATTCCCAAAAGCGAT phtE.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTTGGATGGGA HAJORITY 610 620 630 640 650 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtA.SEO 592 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtB.seo 593 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtB.seo 649 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtB.SEO 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA phtB.SEO AGCAAATGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACTT MAJORITY 660 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGGATAACACTT phtA.SEO 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.seo 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.seo	492 495	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT	C	A TATCATAGA A TATCATTGA A TATCATTGA A TATTATCGA	G G A T A C T G G T G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ
CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtB.seq 545 CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtD.SEO 599 CTTATATCGTTCCTCATGGAGGTCACTATCACTACATTCCCAAAAGCGAT phtE.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTTGGATGGGA Hajority 610 620 630 640 650 658 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtA.SEO 592 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtB.seo 593 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtB.seo 649 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCATCTGGAATGGGA phtB.SEO 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA phtB.SEO AGCAAATGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACCTT Hajority 660 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGGATAACACTT phtA.SEO 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.seo 641 AGCA	492 495	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC	C	ATATCATAGA ATATCATTGA ATATCATTGA ATATTATCGA	GGATACTGGT GGACACGGGT GGACACGGGT AGATACGGGT	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority
CTTATATCGTTCCTCACGGCGACCATTACCATTACATTCCTAAGAATGAG phtD.SEO CTTATATCGTTCCTCATGGAGGTCACTATCACTACATTCCCAAAAGCGAT phtE.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTTGGATGGA Hajority 610 620 630 640 650 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtA.SEO 592 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtB.seo 593 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtB.seo 594 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtB.seo 649 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCATCTGGAATGGGA phtB.SEO 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA phtB.SEO 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGGATAACACTT phtA.SEO 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.seo 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.seo	492 495 549	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC 560	C A A T G C A T C T G J C A A T G C A T C T G J T A A T C C A G C T G J A T G G C G A T C A T 1 570	ATATCATAGA ATATCATTGA ATATCATTGA ATATTATCGA FACCATTACA 580	GGATACTGGT GGACACGGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.sec GATG phtD.SEC AATG phtE.SEC TGAG Majority 600
TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTTGGATGGGA HAJORITY 610 620 630 640 650 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phth.SED 592 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phth.SED 593 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phth.SED 649 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phth.SED 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA phth.SED 640 670 680 690 700 708 AAATCTGTCAAAATTCAAGAACCTATCGCCGACAAAATAGCGGATAACACTT phth.SED 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phth.SED 641 AGCA	492 495 549 608	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC	C A A T O C A T C T O J C A A T O C A T C T O J T A A T C C A G C T C J A T O O C O A T C A T 1 570 A T O O A O A T C A T 1	ATATCATAGA ATATCATTGA ATATCATTGA ATATTATCGA FACCATTACA 580 FACCATTACA	GGATACTGGT GGACACGGGT GGACACGGGT AGATACGGGT TTCCTAAGAA 590	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtA.SEO
610 630 630 640 650 658 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phta.seo 592 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTCCTATCTGGAATGGGA phta.seo 595 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phta.seo 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA phta.seo 640 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT phta.seo 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACACTT phta.seo	492 495 549 608 542 545	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	C A A T G C A T C T G J C A A T G C A T C T G J T A A T C C A G C T G J A T G G C G A T C A T 1 570 A T G G A G A T C A T 1 A C G G C G A C C A T 1 A C G G C G A C C A T 1	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATTATCGA FACCATTACA FACCATTACA FACCATTACA	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA 590 TTCCTAAGAA TTCCTAAGAA	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majorit; 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo
610 630 630 640 650 658 TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phta.seo 592 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTTCCTATCTGGAATGGGA phta.seo 595 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phta.seo 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA phta.seo 640 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT phta.seo 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACACTT phta.seo	492 495 549 608 542 545	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	C A A T G C A T C T G J C A A T G C A T C T G J T A A T C C A G C T G J A T G G C G A T C A T 1 570 A T G G A G A T C A T 1 A C G G C G A C C A T 1 A C G G C G A C C A T 1	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATTATCGA FACCATTACA FACCATTACA FACCATTACA	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA 590 TTCCTAAGAA TTCCTAAGAA	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majorit; 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo
TTATCAGCTAGCGAGTTGGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtA.SEO TTATCAGCTAGCGAGTTGGCTGCTGCAGAAGCCTTCCTATCTGGTCGAGG phtA.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA phtB.SEO TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCATCTGGAATGGGA phtB.SEO AGCAAATGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACTT MAjority 660 670 680 690 700 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT phtA.SEO 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATATG phtB.seo	492 495 549 608 542 545	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	C A A T G C A T C T G J C A A T G C A T C T G J T A A T C C A G C T G J A T G G C G A T C A T 1 A T G G A G A T C A T 1 A C G G C G A C C A T 1 A T G G A G G T C A C 1 A T G G A G G T C A C 1	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTACA S80 FACCATTACA FACCATTACA FACCATTACA FACCATTACA	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA S90 TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEQ TGAG phtB.seq TGAG phtB.seq TGAG phtB.seq CGAT phtE.SEQ
TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA PhtB.sec 595 TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA PhtB.SEC 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA PhtB.SEC AGCAAATGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACTT Majority 660 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT PhtA.SEC 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG PhtB.sec 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG PhtB.sec	492 495 549 608 542 545	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	C A A T O C A T C T O J C A A T O C C A T C T O J T A A T C C A G C T C A T 1 570 A T O O A G A T C A T 1 A C O O C O A C C A T 1 A C O O C O A C C A T 1 A T O O A G O C C A T 1 A T O O A G O T C A C 1 T T A O C T O C T O C C	TATCATAGA ATATCATAGA ATATCATTGA ATATTATCGA FACCATTACA 580 FACCATTACA FACCATTACA FACCATTACA FACCATTACA FACCATTACA	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA S90 TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAT phtE.SEO
TTATCAGCTAGCGAGTTAGCTGCTGCAGAAGCCTATTGGAATGGGA PhtD.SEQ 649 TTATCTGCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAA PhtB.SEQ AGCAAATGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACTT Hajority 660 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT PhtA.SEQ 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG PhtB.seq 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG PhtB.seq	492 495 549 608 542 545 599	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	CAATOCATCTO; CAATOCATCTO; CAATOCATCTO; ATOGCGATCAT1 570 ATOGAGATCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ATGGAGTCAT1	TATCATAGA ATATCATAGA ATATCATTGA ATATTATCGA S80 FACCATTACA FACCATTACA FACCATTACA FACCATTACA FACCATTACA FACCATTACA FACCATTACA FACCATTACA FACCATTACA	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA S90 TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAT phtE.SEO GGGA Majority 650
AGCAAATGGGATCTCGTCCTTCTTCAAGTTCTAGTTATACTT Majority 660 670 680 690 700 708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT phtA.SEp 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.sep 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.sep	492 495 549 608 542 545 599	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG F10 TTATCAGCTAGCGAG	C A A T O C A T C T O J C A A T O C C A T C T O J T A A T C C A G C T C A T 1 570 A T O O C O A T C A T 1 A C O O C O A C C A T 1 A C O O C O A C C A T 1 A C O O C O A C C A T 1 A T O O A G O T C A C 1 T T A O C T O C T O C T	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATTATCGA 580 FACCATTACA AGAAGCC	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA S90 TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA -TATTTGGAT 640	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAT phtE.SEO GGGA Majority 650 GAGG phtA.SEO
660 670 680 690 700 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT phtA.SEp 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.sep 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.sep	608 542 545 599	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG	CAATOCATCTO; CAATOCATCTO; CAATOCATCTO; TAATCCAGCTO; ATGGCGATCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ATGGAGGTCAC1 TAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC;	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTACA 580 FACCATTACA	GGATACTGGT GGACACGGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo GGAT phtE.SEO GGGA Majority 650 GAGG phtA.SEO GGAG phtA.SEO
660 670 680 690 700 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT phtA.SEp 638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.sep 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.sep	492 495 549 608 542 545 599	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATGTCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG	CAATOCATCTO; CAATOCATCTO; TAATCCAGCTO; ATGGCGATCAT; ATGGCGATCAT; ACGGCGACCAT; ACGGCGACCAT; ATGGAGTCAC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC;	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTACA S80 FACCATTACA FACCATT	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo GGAA Majority 650 GGGA Majority 650 GGAG phtA.SEO GGAA phtB.seo GGGAA phtB.seo
708 AAATCTGTCAAATTCAAGAACCTATCGCCGACAAAATAGCGATAACACTT phtA.SEp 638 AG	492 495 549 608 542 545 599	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATATCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG	CAATGCATCTGACATGACATGCATCATACAGCATCATACAGCATCATACAGCATCATACAGCACCATACAGCACCATACAGCACCATACAGCACCATACAGCAGCACCATACAGCAGCACCATACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTACA S80 FACCATTACA FACCATT	GGATACTGGT GGACACGGGT AGATACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtD.SEO AATG phtD.SEO TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo GGAT phtE.SEO GGA Majority 650 GGAG phtA.SEO GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo
638 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtB.eeq 641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtD.SEq	492 495 549 608 542 545 599	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATATCTT CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG	CAATOCATCTO; CAATOCATCTO; CAATOCATCTO; TAATCCAGCTO; ATGGCGATCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ATGGAGGTCAC1 TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC;	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTGA ATATCATTACA 580 FACCATTACA FACCATTCAA CTTCTTCAA CTTCTCAA CTTCTTCAA CTTCTTCAA CTTCTTCAA CTTCTTCAA CTTCTTCAA CTTCTCAA CTTCTTCAA CTTC	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtB.seo GATG phtB.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAA Majority 650 GGGA Majority 650 GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo
641 AGCAGGGATCTCGTCCTTCTTCAAGTTCTAGTTATAATG phtD.SEQ	492 495 549 608 542 545 599	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATATCTT CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG	CAATOCATCTO; CAATOCATCTO; CAATOCATCTO; TAATCCAGCTO; ATGGCGATCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ACGGCGACCAT1 ATGGAGGTCAC1 TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC;	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTGA ATATCATTACA 580 FACCATTACA FACCATTCAA CTTCTTCAA CTTCTCAA CTTCTTCAA CTTCTTCAA CTTCTTCAA CTTCTTCAA CTTCTTCAA CTTCTCAA CTTCTTCAA CTTC	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtB.seo GATG phtB.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAA Majority 650 GGGA Majority 650 GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo
	492 495 549 608 542 545 599 658 592 595 649	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATATCTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG AAATCTGTCAAATTC	CAATOCATCTO; CAATOCATCTO; CAATOCATCTO; TAATCCAGCTO; ATGGCGATCAT? ACGGCGACCAT? ACGGCGACCAT? ACGGCGACCAT? ATGGAGTCAC? TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; GGGATCTCGTCG;	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTGA ATATCATTGA 580 FACCATTACA FACCATTCAAG 680 GCCGACAAAA	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtB.seo GATG phtD.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAA Majority 650 GAGG phtA.SEO GGGA PhtB.seo
695 AAAATATGCAACCGAGTCAGTTA-AGCTATTCTT Phts.SEQ	492 495 549 608 542 545 599 658 592 595 649	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATATCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG AAATCTGCTAAATTCAGCTAGCT	CAATOCATCTO; CAATOCATCTO; CAATOCATCTO; TAATCCAGCTO; ATGGCGATCAT? ACGGCGACCAT? ACGGCGACCAT? ACGGCGACCAT? ATGGAGGTCAC? TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; GGGATCTCGTCG	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTGA ATATCATTACA S80 FACCATTACA GAAGCCTTC AGAAGCCTTC AGAAGCCTTC AGAAGCCTTC AGAAGCCTTC AGAAGCACAAAA CTTCTTCAAG	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtB.seo GATG phtB.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAT phtE.SEO GGAA Majority 650 GGAG phtA.SEO GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo ACTT Majority 700 ACTT phtA.SEO AATG phtB.seo
	492 495 549 608 542 545 599 658 592 595 649	TGATGGGTATATCTT TGATGGTTATATCTT TGATGGTTATATCTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG AAATCTGCTAAATTCAGCTAGCT	CAATOCATCTO; CAATOCATCTO; CAATOCATCTO; TAATCCAGCTO; ATGGCGATCAT? ACGGCGACCAT? ACGGCGACCAT? ACGGCGACCAT? ATGGAGGTCAC? TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; TTAGCTGCTGC; GGGATCTCGTCG	TATCATAGA ATATCATAGA ATATCATTGA ATATCATTGA ATATCATTGA ATATCATTACA S80 FACCATTACA GAAGCCTTC AGAAGCCTTC AGAAGCCTTC AGAAGCCTTC AGAAGCCTTC AGAAGCACAAAA CTTCTTCAAG	GGATACTGGT GGACACGGGT AGATACGGGT TTCCTAAGAA	GATG phtB.seo GATG phtB.seo GATG phtB.SEO AATG phtE.SEO TGAG Majority 600 TGAG phtB.seo TGAG phtB.seo TGAG phtB.seo GGAT phtE.SEO GGAA Majority 650 GGAG phtA.SEO GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo GGGA phtB.seo ACTT Majority 700 ACTT phtA.SEO AATG phtB.seo

10/17

Figure 7(c)

	CAA-ATCCAGCTCAGTA	CCXX	GATTGTCAG	A G A A C C A C A A T	C T Majority
	710	720	730	740	750
758	CAAGAACAAACTGGGT	CCTTCTGTA	AGCAATCCA	GAACTACAAA	TACT phtA.SEQ
677	CAA-ATCCAGCTCAA	. C C X X	GATTGTCAG	A G A A C C A C A A T	C T phtB.seq
680		. •		AGAACCACAAT	
728	CAACAGCT-AGT-		GACAAT A	. C C G C T	CTGT phtE.SEQ
	GACA-AAGCTGTCACTG	CAACATTAT	CA-TCAAGC	AAATCAAGGTG	AAAA Majority
	760	770	780	790	800
808	AACACAAGCAACAACA				
717 720				-	
759	AGCAAAAG-GATCA				
	CATTTCAAGTCTTTTG	GTGAATTGT	ATECTAAAC	CTTTATCAGAA	CCC HAJORITY
	810	820	830	840	850
858	CATTGATAGTCTCTTG	AACAGCTCT	ACAAACTGC	CTTTGAGTCAJ	CGAC phea.seq
756	CATTTCAAGCCTTTTAG	CTGAATTGT	ATGCTAAAC	CCTTATCAGAJ	CGCC phtB.seq
759	CATTTCAAGCCTTTTA				
801	TCTCCAGAGTCTTTTG	CAUGNACICI	ATOATTCAC.		C C I I pace.seq
	ATGTGGAATCTGATGG		GACCCAGCG	CAAATCACAA	TCGA Majority
	860	870	880	890	900
	ATGTAGAATCTGATGG		GATCCAGCA	CALATCACAA	T C G A phtA.SEO
908 806	ATGTEGAATCTGATGG		GACCCAGCG	CAAATCACAA	GTCGA phtB.seq
809	ATGTGGAATCTGATGG	CTTATTTTC	GACCCAGCG	CAAATCACAA	TCGA phtD.SEQ
851	ACAGTGAATCAGATGG	CCTGGTCTTT	GACCCTGCT	A A G A T T A T C A (GTCGT phtE.SEQ
	ACCGCCAGAGGTGTTG			******	A T C C C Majority
	YEEGECX GX GG TO TTO				•
	910	920	930	940	950
958	ACAGCTAGAGGTGTTG	CAGTGCCACA	CGGAGATCA	TTACCACTTC	ATCCC phtA.SEQ
856	ACCGCCAGAGGTGTAG	TGTCCCTCA	TOGTAACCA	TTACCACTTT.	ATCCC pheb.seq
859 901	ACCCCCAGAGGTGTAG		TGGCGACCA		
301	ACACCARA.CO.C.				-
	TTATGAACAAATGTCT	BAATTGGAAG	AACGAATTG	CTCGTATTAT	TCCCC Majority
	960	970	980	990	1000
1006	TTACTCTCAAATGTCT	BAATTGGAAG	AACGAATCG	CTCGTATTAT	TCCCC phtA.SEQ
906	TTATGAACAAATGTCT	AAATTGGAAA	AACGAATTG	CTCGTATTAT	TCCCC phtB.seq
909	· T T A T G A A C A A A T G T C T :	GAATTGGAAA	AACGAATTG	CTCGTATTAT	TCCCC phtD.SEQ
951	TTACAGCAAGCTTTCT	BCCTTAGAAG	DITADAAAA	CCAGAAT	pht2.SEO
	TTCGTTATCGTTCAAA	CCATTOGGTA	CCAGATTCA	AGACCAGAAG	A A C C A Majority
	1010	1020	1030	1040	1050
1051 936	TTCGTTATCGTTCAAA	C C A T T G G G T A		YGGCCYGYYC	AACCA phen.seo
959	TTCGTTATCGTTCAAA	CCATTGGGTA		AGACCAGAAC	AACCA phtD.SEQ
993		GGTG	CCT	ATCAGTGG	AACTG phtE.SEO

11/17

Figure 7(d)

AGTCCACAATCGACTC	CGGAACCTAGTC	CAAGTCCOCA	ACCTGCACCA	A A Hajority
1060	1070	1080	1090	1100
1108 AGTCCACAACCGACTC	CGGAACCTAGTC	CAGGCCCGCA	ACCTGCACCA	A A phtA.SEC
1006 AGTCCACAACCGACTC				
1009 AGTCCACAATCGACTC	CGGAACCTAGTC	CAAGTCCGCA	ACCTGCACCA	A A phtD.SEQ
1013 GTTCTACAGTT	TCTA	C Y Y Y	TGCAA	A A phtE.SEQ
TC-T-AAAGCTCCA	AGCAATCCAATT	G A T G - G A A A T	TGGTCAAAGA	A G Majority
1110	1120	1130	1140	1150
				
1158 TCTTAAAATAGACTCA		GATGGGAAAT	TGGTT AGTC A	A G phtB.seq
••••	AGCAATCCAATT		TOGTCAAAGA	
1039 C C	T	AATG		phtE.SEQ
•				
CTGTTCGAAAAGTAGG	CGATGGTTATGT	CTTTGAGGAG	AATGGAGTTT	CT MAJORITY
1160	1170	1180	1190	1200
1196 TGGTACGAAAAGTTGG				
1088 CTGTTCGAAAAGTAGG				
1109 CTGTTCGAAAAGTAGG	C G A T G G T T A T G T		AATGGAGTTT	
1046 A R G T A G -		C *		yate.and
CGTTATATCCCAGCCA	AGGATCTTTCAG	CAGAAACAGC	AGCAGGCATT	G A Majority
1210	1220	1230	1240	1250
1246 CGTTATGTCTTTGCGA	AAGATTTACCAT	CTGAAACTGT	TAAAAATCTT	G A phtA.SEQ
1138 CGTTATATCCCAGCCA				
1159 CGTTATATCCCAGCCA	AGGÄTETTTCAG	CAGAAACAGC	AGCAGGCATT	G A phtD.SEQ
1062			x G G C	phtB.SEQ
TAGCAAACTGGCCAAG				
. ROURNATION	C	TCTCATAAGC	TAGGAGCTAA	G A Majority
1000		1		
12,60	1270	1280 -	1290	1300
1296 AAGCAAGTTATCAAAA	_ 1270 C A A G A G A G T G T T	1280 - TCACACACTT	1290 TAACTGCTAA	1300 A A phta.seq
1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G	1270 CAAGAGAGTGTT CAGGAAAGTTTA	1280 - TCACACACTT TCTCATAAG	1290 TAACTGCTAA TAGGAACTAA	1300 A A phtA.SEQ G A phtB.seq
1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G	1270 CAAGAGAGTGTT CAGGAAAGTTTA	1280 - TCACACACTT TCTCATAAGC	1290 TAACTGCTAA TAGGAACTAA	1300 A A phtA.SEQ G A phtB.seq
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG	L 1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTA	TCACACACTT TCTCATAAGG TCTCATAAGG	1290 TAACTGCTAA TAGGAACTAA	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtE.SEQ
1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G	L 1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTA	TCACACACTT TCTCATAAGG TCTCATAAGG	1290 TAACTGCTAA TAGGAACTAA	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtE.SEQ
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG	L 1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTA	TCACACACTT TCTCATAAGG TCTCATAAGG	1290 TAACTGCTAA TAGGAACTAA	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtE.SEQ
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TCTCATAAGC	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtE.SEQ G A C Majority 1350
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066 AAACTGATCTTCCTTC 1310 1346 AAGAAAATGTTGCTCC 1238 AAACTGACCTCCCATC	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TC ATTTTACGAT 1330 ATTTTACAAT	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtZ.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066 AAACTGATCTTCCTTC 1310 1346 AAGAAAATGTTGCTCC 1238 AAACTGACCTCCCATC	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TC ATTTTACGAT 1330 ATTTTACAAT	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtZ.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq G A C phtD.SEQ
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066 AAACTGATCTTCCTTC 1310 1346 AAGAAAATGTTGCTCC 1238 AAACTGACCTCCCATC	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TC ATTTTACGAT 1330 ATTTTACAAT	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtZ.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq
1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1066	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TC ATTTTACGAT 1330 ATTTTATGAT ATTTTACAATTTAACGAC	TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG 1340 AAAGCATATA AAGGCTTATG	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtE.SEQ A C Majority 1350 A T phtA.SEQ G A C phtB.seq G A C phtB.seq G A C phtB.seq
1296 A A G C A A G T T A T C A A A A 1188 T A G C A A A C T G G C C A A G 1209 T A G C A A A C T G G C C A A G 1066	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA TAGTGATCGAGA	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TCTCATAAGC TCTCATAAGC ATTTTACGAT ATTTTACAAT ATTTTACAAT ATTTTACAAT	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG AAAGCATATA AAAGCATATA AAAGCTTATG	1300 A A phtA.SEQ G A phtB.seq O A phtD.SEQ phtZ.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq G A C phtB.seq
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066 AAACTGATCTTCCTTC 1310 1346 AAGAAAATGTTGCTCC 1238 AAACTGACCTCCATC 1259 AAACTGACCTCCCATC 1074 - AAGCAATCCTTCTTC TTACTAGCAAGAATTC	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA TAGTGATCGAGA TAGTGATCGAGA	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TCTCATAAGC ATTTTACGAT 1330 ATTTTACAAT ATTTTACAATTTAACGAC TTGATAATAA	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG AAAGCATATA AAAGCATATA AAAGCTTATG	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtE.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq A C phtB.seq A C phtB.seq A C phtD.SEQ phtE.SEQ A C T Majority
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA TAGTGATTTAC	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TCTCATAAGC ATTTTACGAT 1330 ATTTTACGAT ATTTTACAAT ATTTTACAAT TTTACAAT TTGATAATAA	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG AAAGCATATA AAAGCATATA AAAGCTTATG	1300 A A phtA.SEQ G A phtB.seq G A phtD.SEQ phtZ.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq G T Majority 1400 T T C phtA.SEQ
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066 AAACTGATCTTCCTTC 1310 1346 AAGAAAATGTTGCTCC 1238 AAACTGACCTCCATC 1259 AAACTGACCTCCCATC 1074 -AAGCAATCCTTCTTC TTACTAGCAAGAATTC 1360 1396 CTGTTAACTGAGGCTC 1288 TTACTAGCAAGAATTC	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA TAGTGATTTAC ACCAAGATTTAC	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TCTCATAAGC ATTTTACGAT ATTTTACGAT ATTTTACAAT ATTTTACAATTTAACGAC TTGATAATAJ	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG AAGGCTTATG AAGGCTTATG AAGGCTTATG AAGGCTTATG	1300 A A phtA.SEQ G A phtB.seq O A phtD.SEQ phtZ.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq G T Majority 1400 C T C phtA.SEQ G T phtB.seq
1296 AAGCAAGTTATCAAAA 1188 TAGCAAACTGGCCAAG 1209 TAGCAAACTGGCCAAG 1066	1270 CAAGAGAGTGTT CAGGAAAGTTTA CAGGAAAGTTTAAGTCTT TAGTGATCGAGA 1320 TCGTGACCAAGA TAGTGATCGAGA TAGTGATTTAC ACCAAGATTTAC	1280 TCACACACTT TCTCATAAGC TCTCATAAGC TCTCATAAGC ATTTTACGAT ATTTTACGAT ATTTTACAAT ATTTTACAATTTAACGAC TTGATAATAJ TTGATAATAJ TTGATAATAJ	1290 TAACTGCTAA TAGGAACTAA TAGGAACTAA AAGGCTTATG AAGGCTTATG AAGGCTTATG AAGGCTTATG AAGGCTTATG	1300 A A phtA.SEQ G A phtB.seq O A phtD.SEQ phtZ.SEQ A C Majority 1350 A T phtA.SEQ A C phtB.seq G T Majority 1400 C T C phtA.SEQ G T phtB.seq

12/17

PCT/US99/30390

Figure 7(e)

TGATTTTGAG	GCTT' GAT	AACCTGTTGG	BAACGACTCAA	GGATGTCTCAA	Majority
141	0 1	420	1430	1440 145	50
1446 TGATTTCCAA	GCCTTAGAC	AAATTATTAG	AACGCTTGAA	TGATGAATCGA	phtA.SZO
1338 T G A T T T T G A G	GCTTTGGAT	AACCTGTTGG	. A A C G A C T C A A		
1359 TGATTTGAG	GCTTTGGAT	AACCTGTTGG		A G G A T G T C C C A A	•
1107	• • • • • • • • • • • • • • • • • • •		• • • • • • • • • •	• • • • • • • • • • • • • • • • • • • •	phtE.SEQ
GTGATAAAGT	CAAGTTAGT	GGATGATATT		TAGCTCCGATT	Majority
146				1490 150	_
1496 CTAATAAAGA	AAAATTGGT	AGATGATTTA			_
1388 GTGATAAAGT					
1409 GTGATAAAGT		GGATGATATT			
1107			0 CT CT CT	7	phtE.SEQ
COTCATCCAG	. A C G T T T A G	G A A A A C C A A A	LTGCGCAAAT1	. A C C T A C A C T G A	Majority
1510					•
		1			_
1546 ACCCATCCAG. 1438 CGTCATCCAG.					
1459 CGTCATCCAG					
1115					phtE.SEC
				. 	
TGATGAGATT				ZAGCATCAGATO	WYJOLIEA
1560		1		1590 16	_
1596 AGACGAAGTT					
					•
1488 TGATGAGATT	CAAGTAGCC	AAGTTGGCAG	GCAAGTACAC		phtB.sec
	CAAGTAGCC	AAGTTGGCAG	G C A A G T A C A C		phtB.sec phtD.SEC
1488 TGATGAGATT	CAAGTAGCC	AAGTTGGCAG	G C A A G T A C A C	:	phtB.sec phtD.SEQ
1488 T G A T G A G A T T 1 1509 T G A T G A G A T T 1 1115	C	AAGTTGGCAG	G C	:	phtB.sec phtD.SEC phtE.SEC
1488 T G A T G A G A T T 1 1509 T G A T G A G A T T 1 1115	C	AAGTTGGCAG AAGTTGGCAG TGATATAACC	GCAAGTACAC GCAAGTACAC GCAAGTACAC	: A G C A G A A G A C G : A A C A G A A G A C G : A G C A T C T G A T G	phtB.sec phtD.SEC phtE.SEC
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T 1115	C A A G T A G C C C A A G T A G C C C C C C C C C C C C C C C C C C	AAGTTGGCAG AAGTTGGCAG TGATATAACC	G C A A G T A C A C G C A A G T A C A C C A G T G A T G A G G	2 A G C A G A A G A C G 2 A A C A G A C A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 1640 16	phtB.seo phtD.SEO phtE.SEO Majority
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T T T T T T T T T	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G	AAGTTGGCAGAAGTTGATAACC	3 G C A A G T A C A C A C A C A C A C A C A C A C A	2 A G C A G A A G A C G 2 A A C A G A A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4 G G G A T G C C T A T 5 G G G A T G C C T A T	phtB.seo phtD.SEO phtE.SEO Hajority phtA.SEO phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T T T T T T T T T	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G	AAGTTGGCAGAAGTTGATAACG	3 G C A A G T A C A C A C A C A C A C A C A C A C A	2 A G C A G A A G A C G 2 A A C A G A C A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4640 16	phtB.seo phtD.SEO phtE.SEO Kajority phtA.SEO phtA.SEO phtB.seo phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T T T T T T T T T	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G	AAGTTGGCAGAAGTTGATAACC	3 G C A A G T A C A C A C A C A C A C A C A C A C A	2 A G C A G A A G A C G 2 A A C A G A A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4 G G G A T G C C T A T 5 G G G A T G C C T A T	phtB.seo phtD.SEO phtE.SEO Hajority phtA.SEO phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	CAAGTAGCCCAAGTAGCCCAAGTAGCCCCCCGCGCAATCCCCCGCGTAATCCCCCCCGTAATCCCCCCGTAATCCCCCCGTAATCCCCCCCGTAATCCCCCCC	AAGTTGGCAGAAGTTGATAACC	GCAAGTACAC GCAAGTACAC CAGTGATGAGG L630 CAGTGATGAG CAGTGATGAG	2 A G C A G A A G A C G 2 A A C A G A A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4 G G G A T G C C T A T 5 G G G A T G C C T A T	phtB.seo phtD.SEO phtE.SEO Kajority DhtA.SEO phtA.SEO phtB.seo phtB.seo phtD.SEO phtE.SEO
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G	AAGTTGGCAGAAAGTTGATATAACCTGATATAACCTGATATAACC	GCAAGTACAC GCAAGTACAC AGTGATGAGG AGTGATGAGG AGTGATGAGG	CAGCAGAAGACG CAACAGAAGACG CAGCATCTGATG GGGATGCCTAT	phtB.seo phtD.SEO phtE.SEO Hajority phtA.SEO phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T T 1115	CAAGTAGCCCAAGTAGCCCTCGCTGATCCTCGTGATCCTCGTGATCCTCGTAATCCTCG	AAGTTGGCAGAAGTTGATAACC	3 G C A A G T A C A C S G C A A G T A C A C C C C C C C C C C C C C C C C	2 A G C A G A A G A C G 2 A A C A G A A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4 G G G A T G C C T A T 5 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G A T A G T T T G T C 6 G G G T A G T T T G T C 6 G G G G T G C T A T 6 G G A T A G T T T G T C 6 G G G G G T G C T A T 6 G G G G G T G C C T A T 6 G G G G G T G C C T A T 6 G G G G G T G C C T A T 6 G G G G G G G G G G G G G G G G G G G	phtB.seo phtD.SEO phtE.SEO Hajority phtA.SEO phtB.seo phtB.seo phtD.SEO phtB.seo htD.SEO
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G T A A T C C T C G A T A T G A C C C	AAGTTGGCAGAAAGTTGATAAACCTGATATAAACCTGATATAACC	3 G C A A G T A C A C 3 G C A A G T A C A C 3 A G T G A T G A G G 3 A G T G A T G A G G 3 A G T G A T G A G G 4 A G T G A T G A G G 5 A G T G A T G A G G 6 A G T G A T G A A A A 1680 6 G A T T G G A A A	2 A G C A G A A G A C G 2 A A C A G A A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4 G G G A T G C C T A T 5 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G G A T G C C T A T 6 G G A T A G T T T G T C 6 G G G T A G T T T G T C 6 G G G G T G C T A T 6 G G A T A G T T T G T C 6 G G G G G T G C T A T 6 G G G G G T G C C T A T 6 G G G G G T G C C T A T 6 G G G G G T G C C T A T 6 G G G G G G G G G G G G G G G G G G G	phtB.seo phtD.SEO phtE.SEO Hajority phtA.SEO phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G A T A T G A C C C A T A T G A C C C A T A T G A C C C	AAGTTGGCAGAAAGTTGATAAACCTGATATAAACCTGATATAACC	1630 3 G A T T A A A A A A A A A A A A A A A A	2 A G C A G A A G A C G 2 A A C A G A A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4 G G G A T G C C T A T 5 G G G A T G C C T A T 6 G G G A T G C C T A T 7 G G G A T G C C T A T 7 G G G A T G C C T A T 7 G G G A T G C C T T T G 8 G G A T A G C T T T G 8 G G A T A G C T T T G 8 G G A T A G C T T T G 8 G A T A G T T T G T C	phtB.seo phtD.SEO phtE.SEO Hajority phtA.SEO phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G A T A T G A C C C A T A T G A C C C A T A T G A C C C	AAGTTGGCAGAAAGTTGATAAACCTGATATAAACCTGATATAACC	1630 3 G A T T A A A A A A A A A A A A A A A A	2 A G C A G A A G A C G 2 A A C A G A A G A C G 2 A G C A T C T G A T G 3 G G G A T G C C T A T 4 G G G A T G C C T A T 5 G G G A T G C C T A T 6 G G G A T G C C T A T 7 G G G A T G C C T A T 7 G G G A T G C C T A T 7 G G G A T G C C T T T G 8 G G A T A G C T T T G 8 G G A T A G C T T T G 8 G G A T A G C T T T G 8 G A T A G T T T G T C	phtB.seo phtD.SEO phtE.SEO Majority phtA.SEO phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T T 1115	CAAGTAGCCCAAGTAGCCCTCGCTGATCCTCGTAATCCTCGTAATCCTCGTAATCCTCGTAATCCTCGTAATGACCCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACCAATATGACAATATGACCAATATGACCAATATGACCAATATGACAATATGACCAATATGACCAATATGACAATATGACCAATATTGACCAATATATTATATATA	AAGTTGGCAGAAAGTTGGTAAACGTGATATAACGTGATATAACGTGATATAACGTGATATAACGTGATAGGCACTGATAGGCCACTGATAGGCCACTGATAGGCCACTGATAGGCCACTGATAGGCCACTGATAGGCCACTG	GCARGTACAC GCAAGTACAC CAGTGATGAGG AGTGATGAGG AGTGATGAGG AGTGATGAGG AGTGATGAGG AGTGATGAGG AGTGATGAGG AGTGATGAGAGG AGTGATGAGAGG	2 A G C A G A A G A C G 2 A C A G A A G A C G 2 A G C A T G C T A T 3 G G G A T G C C T A T 3 G G G A T G C C T A T 4 G A G A T G C C T A T 4 G A T A G T T T G T G 4 G A T A G C T T T T G T G 4 G A T A G T T T T G T G 4 G A T A G T T T T G T G 5 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T G T G 6 A G A T A G T T T T G T G 6 A G A T A G T T T T G T G 7 A G A T A G T T T G T G 7 A G A T A G T T T G T G 7 A G A T A G T T T G T G 7 A G A T A G T T T G T G 7 A G A T A G T T T G T G 7 A G A T A G T T T G T G 7 A G A T A G T T T G T G 7 A G A T A G T T T G T G 7 A G T T T G T G 7 A G T T T G T G 7 A G T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T G T G 7 A G T T T T T G T G 7 A G T T T T T G T G 7 A G T T T T T T T T T T T T T T T T T T	phtB.seo phtD.SEO phtE.SEO phtA.SEO phtB.seo
1488 T G A T G A G A T T 1509 T G A T G A G A T T T 1115	C A A G T A G C C C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C C A G A G C G G C A	AAGTTGGCAGAAAGTTGATAAACCTGATAATAACCTGATATAACCACTGATATAACCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGCACTGATAGATA	1630 AGTGATGAG AGTGATGAGAG AGTGATTAAAAA AGATTAAAAA	2 A G C A G A A G A C G 2 A C A G A A G A C G 3 A C A G A T C T G A T G 4 G A G A T G C C T A T 4 G A G A T G C C T A T 4 G A T A G T T T G T C 6 A G A T A G T T T G T C 6 A G A T A G T T T G T C 6 A G A T A G T T T G T C 6 A G A T A G T T T G T C 6 A G A T A G G T T T G T C 6 A G A A A G G T T T G T C 6 A G A A A G G T T T G T C	phtB.seo phtD.SEO phtE.SEO Majority phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C G C A A G A G C G G C A 1	AAGTTGGCAGAAAGTTGATAAACCTGATAATAAACCTGATATAAACCTGATATAACCACTGATATAACCACTGATAGCACTGATAGATA	3 G C A A G T A C A C A C A C A C A C A C A C A C A	2 A G C A G A A G A C G 2 A C A G A A G A C G 3 A C A G A T C C T A T 4 G A G A T G C C T A T 5 G G G A T G C C T A T 6 G A G A T G C C T A T 6 G A T A G T T T G T C 6 A G A T A G T T T G T C 6 A G A T A G T T T G T C 6 A G A T A G G T T T G T C 6 A G A T A G G T T T G T C 6 A G A A A G G T T T G T C 6 A G A T C C T A T T C 6 A G A T C C T A T T C T C 6 A G A T C C T A T T C T C 7 A G A T C C T T T C T C 7 A G T T C T C T C T T T C T C 7 A G T T T C T C T T T C T C T T T C T C T	phtB.seo phtD.SEO phtE.SEO Majority phtA.SEO phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo phtB.seo thajority majority
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A C A C C C C C C C C C C C C	A A G T T G G C A G A A G T T G G C A G T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C A T A G C C A C T G	1630 2 A G T G A T G A G 2 A G T G A T G A G 2 A G T G A T G A G 3 A G T G A T G A G 3 A G T G A T G A G 4 A G T G A T G A A A A 5 A G T G A T A A A A A 6 G A T T A A A A A 6 G A T T A A A A A 6 G A T T A A A A A 6 G A T T A A A A A 6 G A T T A A A A A 7 A T G C T A A A G 7 A T A C T A A A C T A A A G 7 A T A C T A A A C T A A A C T A C T A	2 A G C A G A A G A C G 2 A C A G A A G A C G 3 A C A G A T C C T A T 4 G A G A T G C C T A T 5 G G G A T G C C T A T 6 G A G A T G C C T A T 7 A G A T A G T T T G T C 8 A G A T A G T T T G T C 8 A G A T A G T T T G T C 8 A G A T A G T T T G T C 8 A G A T A G G T T T G T 8 A G A A A G G T T T G T 8 A G A A A G G T T T C T 8 A G A A A G G T A T C C	phtB.seo phtD.SEO phtE.SEO Majority phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C C A T A T G A C C C C C A C A C C C C C C C C C C C	A A G T T G G C A G A A G T T G G C A G T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C A T A G C C A C T G A T A G C C A C T G A T A G C C A C T G A T A G C C A C T G G C C C A G C T T G C C C A G C C T G G C C C A G C C T G G C C C A G G C T T G C C C A G G C T T	1630 AGTGATGAG AGTGATGAAA AGTGATAAAAA ATGCTAAAG ATGCTAAAG ATGCTAAAG	2 A G C A G A A G A C G 2 A C A G A A G A C G 3 A C A G A T C C T A T 4 G A G A T G C C T A T 5 G G G A T G C C T A T 6 G G G A T G C C T A T 7 G G G A T G C C T A T 7 G G G A T G C C T A T 8 G G G A T G C C T A T 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A A A G G T T T G T 8 G A T T T G T 8 G A A A G G T T T G T 8 G A A A G G T T T G T 8 G A A A G G T T T G T 8 G A A A G G T T T G T 8 G A T T T G T 8 G A T T T T T T T T T T T T T T T T T T	phtB.seo phtD.SEO phtE.SEO Majority phtB.seo
1488 T G A T G A G A T T T 1509 T G A T G A G A T T T T 1115	C A A G T A G C C C A A G T A G C C C A A G T A G C C T G A T C C T C G T G A T C C T C G T G A T C C T C G T G A T C C T C G T A A T C C T C G A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C A T A T G A C C C C C A C A C C C C C C C A C A C	A A G T T G G C A G A A G T T G G C A G T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C T G A T A T A A C C A T A G C C A C T G A T A G C C A C T G A T A G C C A C T G A T A G C C A C T G G C C C A G C T T G C C C A G C C T G G C C C A G C C T G G C C C A G G C T T G C C C A G G C T T	1630 AGTGATGAG AGTGATGAAA AGTGATAAAAA ATGCTAAAG ATGCTAAAG ATGCTAAAG	2 A G C A G A A G A C G 2 A C A G A A G A C G 3 A C A G A T C C T A T 4 G A G A T G C C T A T 5 G G G A T G C C T A T 6 G G G A T G C C T A T 7 G G G A T G C C T A T 7 G G G A T G C C T A T 8 G G G A T G C C T A T 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A T A G T T T G T C 8 G A A A G G T T T G T 8 G A T T T G T 8 G A A A G G T T T G T 8 G A A A G G T T T G T 8 G A A A G G T T T G T 8 G A A A G G T T T G T 8 G A T T T G T 8 G A T T T T T T T T T T T T T T T T T T	phtB.seo phtD.SEO phtE.SEO Majority phtB.seo

13/17

Figure 7(f)

CCCCTCCTTCGACAG	CATCAGGATTCAGGAAATA	CTGAGGCAAAAGGA Hajority
1760	1770 1780	1790 1800
1706 7 4 5 5 7 5 5 7 5 7 5 7 5 7 5 7 5 7 5 7	CGCAGATGTTAAAGCAAATC	CAACTGGAGATAGT obta SEO
	CCATCAGGATTCAGGAAATA	
	CCATCAGGATTCAGGAAATA	
1167		phtE.SEQ
GCAGAAGCTATCTACA	ACCGXGTGAAAGCAGCTAAG	AAGGTGCCACTTGA Hajority
1810	1820 1830	1840 1850
1846 GCAGCAGCTATTACA	ATCGTGTGAAAGGGGAAAAA	CGAATTCCACTCGT phtA.SEQ
.1738 GCAGAAGCTATCTACA	ACCGHGTGAAAGCAGCTAAG.	AAGGTGCCACTTGA phtB.seq
	ACCGCGTGAAAGCAGCTAAG	AAGGTGCCACTTGA phtD.SEQ
1167 TACA	GCT	phtE.SEQ
TCCT1TCCTT1C11T	CTTCAATATACTGTAGAAGT	Class CCCTlCTT Weignity
		- T
1860	1870 1880	1890 1900
	GTTGAGCATACAGTTGAGGT	
	CTTCAATATACTGTAGAAGT	
	CTTCAATATACTGTAGAAGT	
1174	TATATIUTAAGA	phts.seq
TAATCATACCTCATTA	TGATCATTACCATAACATTA	AATTTGAGTGGTTT Majority
1910	1920 1930	1940 1950
1946 7 5 1 7 7 1 7 7 7 7 7 7 7 7 7 7 7 7 7 7	GGATCATTACCATAATATTA	
	TGACCATTACCATAACATCA	
	TGACCATTACCATAACATCA	
1186 C A T G G	TGATCATTTCCATTACATT -	phtE.SEC
GACGAAGGEETTTATO	AGGCACCTAAGGGGTATACT	CTTGAGGATCTTTT MAJORITY
1960	1970 1980	1990 2000
1996 GATGATCACACATACA	AAGCTCCAAATGGCTATACC	TTGGAAGATTTGTT phtA.SEQ
	AGGCACCTAAGGGGTATACT	
	AGGCACCTAAGGGGTATACT	
1210	CCXXX	phtE.SZO
GGCGACTGTCAAGTAC	TATGTCGAACATCCAGACGA	ACGTCCGCATTCAG Majority
2010	2020 2030	2040 2050
	TACGTAGAACACCCTGACGA	· · · · · · · · · · · · · · · · · · ·
	TATGTCGAACATCCAAACGA	
1111	CARATTOOUCKACC-OKC	
ATAATGGTTTTGGTAA	CGCTAGCGACCATGTTTTXA	- A A A C A A G A A A G A T Majority
2060	2070 2080	2090 2100 .
	. T G C C A G T G A G C A T G T G T T . C G C T A G C G A C C A T G T T C A A A	
	. C G C T A G C G A C C A T G T T C A A A	•
	CACCTTCT-CCATCTCTTC-	The state of the s
		· · · · · · · · · · · · · · · · · · ·

14/17

Figure 7(g)

CAAGCCAGTAA	ACCTAATGAAGATG	AGAAACATGACC	AGTAAG-GAG-	- Majority
2110	2120	2130		150
2143 CA	CAGTGAAGAT-			- phtA.SED
2038 C				
2059 C A A G A C A G T A A	ACCTGATGAAGATA		AGTAAGTGAGC	
1284 CAATCCAG - GA	ACTTCACATG;	AGAAACATGA		- phtE.SEQ
A - C T C A C - G	A A T G A A G A A C	3	`*; :	- Madaadaa
2160	2170	2180		T '
2164				200 —
	AAAACCTGAGGAAG		- T T C A A A	
2109 AACTCACCCTG	AATCTGATGAAAAA	3		C pht8.seq
1314	AGAA6	ATGGATACG G	ATTTGA-TGCT	
-1663617333				,
1	CCGTATAAGCCAG-		- A - A C A	A Hajority
2210	2220	2230	2240 2	250
	GAGCCAG			- phtA.SEQ
2114 G A G A A G A G A A A (2159 C A G C A G A T A A T (ACCAGAGTCTCC		
1339		ACTGATACGGAA	GYCYCYCY	•
			•••••	- phtz.seq
G-AGCTGGAGGI	AXCACCAGATGAGT	CAGAAGTXCCTC	AAGTAGAGACT	G Majority
2260	2270	2280		T 300
2222				
SINS T Y G Y C C 1	AACACCTGCTGAGC	CAGAAGTCCCTC	AAGTAGAGACT	4
2163 GGAACCAGAAGJ		CAGAAGAACCTC	AAGTAGAGACT	G phtA.SEQ
2163 GGAACCAGAAGJ 2209 GAAGCTGAAGA1	T A A D D A D A D D A D T A A D D A D T A D A - D A D D A ?		AAGTAGAGACT	G phtA.SEQ
2163 GGAACCAGAAGJ			AAGTAGAGACT	G phtA.SEQ
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A 1 1351 G C T G A A G A -	T A A D B A D A C A C A C A C A C A C A C A C A C	CAGAAGAACCTC CTGAAATTCCTC	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAT	G phtA.SEQ G phtB.seq T phtD.SEQ - phtE.SEQ
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A 1 1351 G C T G A A G A -	T A A D D A D A D D A D T A A D D A D T A D A - D A D D A ?	CAGAAGAACCTCCTCCTGAAAATTCCTC	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAA CTTGXAAAAGT	G phtA.SEQ G phtB.seq T phtD.SEQ - phtE.SEQ C Hajority
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A 1 1351 G C T G A A G A - AAAA X G T T G A A G 2310	2320	CAGAAGAACCTC CTGAAATTCCTC C GCXGAGGTTTTO 2330	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAT 	G phth.SEQ G phth.seq T phtD.SEQ - phtE.SEQ C Majority
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A T 1351 G C T G A A G A - AAA A X G T T G A A G 2310 2214 A A A A A G T A G A A G 2213 A A A A A G G T T G A A G	ATCACCAGAGGAAT ACCAC-AGATGAGG TGAAT BCXAAACTXAXAGAX 2320 BCCCAACTCAAGAA	CAGAAGAACCTC CTGAAATTCCTC C GCXGAGGTTTTG 2330 GCAGAAGTTTTG	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAT CTTGXAAAAGT 2340 2:	G phtA.SEQ G phtB.seq T phtD.SEQ - phtE.SEQ C Majority 150 A phtA.SEQ
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A T 1351 G C T G A A G A C 2310 2214 A A A A A G T A G A A G 2213 A A A A A G G T T G A A G 2258 C T G T T A T T A A C G	ATCACCAGAGGAAT ACCAC-AGATGAGG TGAAT BCXAAACTXAXAGAX 2320 BCCCAACTCAAGAA	CAGAAGAACCTC CTGAAATTCCTC C GCXGAGGTTTTG 2330 GCAGAAGTTTTG	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAT CTTGXAAAAGT 2340 2:	G phtA.SEQ G phtB.seq T phtD.SEQ - phtE.SEQ C Majority 150 A phtA.SEQ
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A T 1351 G C T G A A G A - AAA A X G T T G A A G 2310 2234 A A A A A G T A G A A G	ATCACCAGAGGAAT ACCAC-AGATGAGG TGAAT BCXAAACTXAXAGAX 2320 BCCCAACTCAAGAA	CAGAAGAACCTC CTGAAATTCCTC C GCXGAGGTTTTG 2330 GCAGAAGTTTTG	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAT CTTGXAAAAGT 2340 2: CTTGCGAAAGT CTTGGAAAAAT	G phtA.SEQ G phtB.seq T phtD.SEQ - phtE.SEQ C Majority 150 A phtA.SEQ
2216 AAAAAGTTGAAG 2213 AAAAAGTTGAAG 2214 AAAAAGTTGAAG 2213 AAAAAGTTAAAG 2215 CTGTTATTAACG	ATCACCAGAGGAATACACACACACACACACACACACACAC	CAGAAGAACCTC CTGAAATTCCTC C GCXGAGGTTTTG 2330 GCAGAAGTTTTG GCTGAAGATTTA GCGGAGGCCTTGAGGTTTTG	AAGTAGAGACT AAGTAGAGACT AAGTAGAGAAT CTTGXAAAAGT 2340 2: CTTGCGAAAGT CTTGGAAAAAT	G phth.SEQ G phth.SEQ T phtD.SEQ - phtE.SEQ C Majority 150 A phth.SEQ C phtB.SEQ A phtD.SEQ C phtB.SEQ
2163 G G A A C C A G A G G G A G C A G A G A	ATCACCAGAGGAAT ACCAC-AGATGAGG CCAAACTXAXAGAX 2320 CCCAACTCAAAGAA AAAACTGAGAGAG	CAGAAGAACCTC CTGAAATTCCTC C GCXGAGGTTTTG 2330 GCAGAAGTTTTG GCTGAAGATTTA GCGGAGGCCTTGAGGTTTTG	AAGTAGAGACT AAGTAGAGACT AAGTAGAGAAT CTTGXAAAAGT 2340 2: CTTGCGAAAGT CTTGGAAAAAT	G phth.SEQ G phth.SEQ T phtD.SEQ - phtE.SEQ C Majority 150 A phth.SEQ C phtB.SEQ A phtD.SEQ C phtB.SEQ
2163 G G A A C C A G A G G G A G C A G A G A	ATCACCAGAGGAAT ACCAC-AGATGAGG CCAAACTXAXAGAX 2320 CCCAACTCAAAGAA AAAACTGAGAGAG	CAGAAGAACCTC CTGAAATTCCTC CTGAAGTTTTG 2330 GCAGAAGTTTTG GCTGAAGATTTTA GCGGAGGCCTTGAGGTTTTG XACGGAGACTCT	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAA CTTGXAAAAGT 2340 2: CTTGCGAAAGT CTTGGAAAAAT CTAGAAAAAGT	G phth.SEQ G phth.seq T phth.SEQ - phtE.SEQ C Hajority 150 A phth.SEQ C phtB.seq A phtD.SEQ C phtB.seq A htD.SEQ C phtE.SEQ
2163 G G A A C C A G A G G G A G C A G A G G A G C T G A A G A T C T G A A G A T C T A G T A G A A G C T G A A G A A G C T G A A G C T G A A G C T G A A G C T G A A G C T G A A G C T G A A G C T G T A T T A A C G C T A G T A G A G C T A G T A G A G C T A G T A G A G C C T A G T C T A G	ATCACCAGAGGAAT ACCACCAGAGGAAT ACCACCAGATGAGG 2320 CCCAACTCAAAGAA AAAACTGAGAGAG CTAAGATAGCAGAT ATXAAAXCCAATGC 2370 CTGAAAGCCAATGC	CAGAAGAACTCT CTGAAATTCCTC CTGAAATTCCTC CCTGAAGGTTTTG 2330 GCAGAAGTTTTG GCTGAAGATTTA GCGGAGGCCTTGAGGTTTTG XACGGAGACTCT 2380 AACAGAAACTCT	AAGTAGAGACT AGGTCGAGACT AAGTAGAGAAA CTTGXAAAAGT 2340 2: CTTGCGAAAGT CTTGGAAAAAT CTAGAAAAAGT XACTGGTTTAA	G phth.SEQ G phth.seq T phth.SEQ - phtE.SEQ C Hajority 150 A phth.SEQ C phtB.seq A phtD.SEQ C phtB.seq A htD.SEQ C phtE.SEQ
2214 AAAAAGTTGAAG 2213 AAAAAGTTGAAG 2214 AAAAAGTTGAAG 2213 AAAAAGTTAAAG 2218 CTGTTATTAACG 1365	ATCACCAGAGGAAT ACCACAGAGGAAT ACCACACTXAXAGAX 2320 CCCAACTCAAAGAA AAAAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	CAGAAGAACTCTCCAGAAAACTCTCCAGAAAATTCCTCCCCCCCC	AAGTAGAGACT AAGTAGAGACT AAGTAGAGACT AAGTAGAGAAAA CTTGXAAAAAGT CTTGCGAAAAAT CTAGAAAAAAT CTAGAAAAAAT CTAGAAAAAAT AACTGGTTTAA 2390 24	G phth.SEQ G phth.seq T phth.SEQ - phtz.SEQ C Hajority ISO A phth.SEQ C phtB.Seq A phtD.SEQ C phtz.SEQ A Majority T C phtz.SEQ A Majority T C phth.SEQ C phtz.SEQ A phth.SEQ
2163 G G A A C C A G A G G G A G C A G A G G A G C T G A A G A T C T G A A G A T C T A G T A G A A G C T G A A G A A G C T G A A G C T G A A G C T G A A G C T G A A G C T G A A G C T G A A G C T G T A T T A A C G C T A G T A G A G C T A G T A G A G C T A G T A G A G C C T A G T C T A G	ATCACCAGAGGAAT ACCACAGAGGAAT ACCACACTXAXAGAX 2320 CCCAACTCAAAGAA AAAAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	CAGAAGAACTT CTGAAATTCCTC CTGAAATTCCTC 2330 CCAGAAGTTTTG CCTGAAGATTTTG CCTGAAGATTTTA CCGGAGGCCTTG AGGTTTTG XACGGAGACTCT 2380 AACAGAAACTCT CAAAGAGACACTTT TATGGAGACACTTT	AAGTAGAGACT AAGTAGAGACT AAGTAGAGACT AAGTAGAGAAAA CTTGXAAAAAGT CTTGCGAAAAAT CTAGAAAAAAT CTAGAAAAAAT CTAGAAAAAAT AACTGGTTTAA 2390 24 AGCTGGTTTAA GACTGGTTTAA	G phth.SEQ G phth.seq T phth.SEQ - phtz.SEQ C Hajority ISSO A phth.SEQ C phtB.Seq A phtD.SEQ C phtz.SEQ A Majority T phth.SEQ A phth.SEQ
2209 GAAGCTGAAGA 2209 GAAGCTGAAGA 1351 GCTGAAGA 2310 2214 AAAAAGTTGAAG 2213 AAAAAGTTGAAG 2213 AAAAGTTGAAG 21365	CCCAACTCAAAGA CTAAACTCAAAGA CTAAACTCAAAGA AAAAACTGAGAGG CTAAGATAGCAGAT ATXAAACCCAATGC 2370 CTGAAAGCCAATGC ATCAAGTCAAAGC	CAGAAGAACTT CTGAAATTCCTC 2330 CCAGAAGTTTTG CCAGAAGATTTTG CCAGAAGATTTTG CCAGAAGATTTTG CCAGAAGATTTA CCGGAGACTCT 2380 AACAGAAACTCT CAAAGAGACTCT TATGGAGACACTT -ACGGAGACTCT	### ### ### ### ### ### ### ### ### ##	G phtA.SEQ G phtB.SEQ T phtD.SEQ - phtE.SEQ C Majority A phtA.SEQ C phtB.SEQ A phtD.SEQ A phtD.SEQ A Majority G phtZ.SEQ A phtB.SEQ
2209 GAAGCTGAAGA 2209 GAAGCTGAAGA 1351 GCTGAAGA 2310 2214 AAAAAGTTGAAG 2213 AAAAAGTTGAAG 2213 AAAAGTTGAAG 21365	ATCACCAGAGGAAT ACCACAGAGGAAT ACCACACTXAXAGAX 2320 CCCAACTCAAAGAA AAAAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	CAGAAGAACTT CTGAAATTCCTC 2330 CCAGAAGTTTTG CCAGAAGATTTTG CCAGAAGATTTTG CCAGAAGATTTA CCGGAGACTTTTG XACGGAGACTCT 2380 AACAGAAACTCT CAAAGAGACTCT TATGGAGACACTTT	### ### ### ### ### ### ### ### ### ##	G phtA.SEQ G phtA.SEQ T phtD.SEQ - phtE.SEQ C Majority A phtA.SEQ C phtB.seq A phtD.SEQ C phtB.seq A phtD.SEQ A phtB.seq A phtD.SEQ A phtB.seq A phtD.SEQ - phtE.SEQ A phtB.seq A phtB.seq A phtB.seq A phtB.seq A phtD.SEQ
2209 GAAGCTGAAGA 2209 GAAGCTGAAGA 1351 GCTGAAGA 2310 2214 AAAAAGTTGAAG 2213 AAAAAGTTGAAG 2213 AAAAGTTGAAG 21365	CCCAACTCAAAGA CTAAACTCAAAGA CTAAACTCAAAGA AAAAACTGAGAGG CTAAGATAGCAGAT ATXAAACCCAATGC 2370 CTGAAAGCCAATGC ATCAAGTCAAAGC	CAGAAGAACTT CTGAAATTCCTC 2330 CCAGAAGTTTTG CCAGAAGATTTTG CCAGAAGATTTTG CCAGAAGATTTA CCGGAGACTTTTG XACGGAGACTCT 2380 AACAGAAACTCT CAAAGAGACTCT TATGGAGACACTTT	AAGTAGAGACT AAGTAGAGACT AAGTAGAGACT AAGTAGAGAAA CTTGXAAAAAGT CTTGCGAAAAAT CTAGAAAAAGT XACTGGTTTAA 2390 24 AGCTGGTTTAA CACAGGATTAA	G phtA.SEQ G phtA.SEQ T phtD.SEQ - phtE.SEQ C Majority A phtA.SEQ C phtB.seq A phtD.SEQ C phtB.seq A phtD.SEQ A phtB.seq A phtD.SEQ A phtB.seq A phtD.SEQ - phtE.SEQ A phtB.seq A phtB.seq A phtB.seq A phtB.seq A phtD.SEQ
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A T 1351 G C T G A A G A T 2310 2234 A A A A A A G T A G A A G 2213 A A A A G G T T G A A G 2258 C T G T T A T T A A C G 1365	ATCACCAGAGGAAT ACCACAGAGGAAT ACCACAACTXAXAGAX 2320 CCCAACTCAAAGAAAAACTGAGAGAGAGAGAGAGAGAGAG	CAGAAGAACTC CTGAAATTCCTC 2330 CCAGAAGTTTTG 2330 CCAGAAGTTTTG CCTGAAGATTTA CCGGAGACTTTA CCGGAGACTTTT 2380 AACAGAAACTCT CAAAGAGACTCT TATGGAGACACTT TATGGAGACACTT AACAGAGACTCT AACAGAGAGACTCT AACAGAGAGAGACTCT AACAGAGAGAGACTCT AACAGAGAGAGACTCT AACAGAGAGAGACTCT AACAGAGAGAGACTCT AACAGAGAGAGACTCT AACAGAGAGAGACT	### CTT G C A G A CT A G T CTT G C A A A A G T CTT G C A A A A G T CTT G C A A A A G T CTT G CTT CTT CTT CTT CTT CTT CTT C	G phtA.SEQ G phtB.seq T phtD.SEQ - phtE.SEQ C Majority TSO A phtA.SEQ C phtB.seq A phtD.SEQ C phtE.SEQ A phtD.SEQ A phtD.SEQ A phtD.SEQ A phtB.seq A phtD.SEQ C phtA.SEQ A phtB.seq A phtD.SEQ C phtB.seq
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A T 1351 G C T G A A G A T 2310 2234 A A A A A A G T T G A A G 2213 A A A A G G T T G A A G 2258 C T G T T A T T A A C G 1365	ATCACCAGAGGAAT ACCACAGAGGAAT ACCACACTXAXAGAX 2320 CCCAACTCAAAGAA AAAAACTGAGAGAGAGAGAGAGAGAGAGAGAAAACTGAGAGAGA	CAGAAGAACCTC CTGAAATTCCTC CTGAAATTCCTC CCTGAAGTTTTG 2330 GCAGAAGTTTTG GCTGAAGATTTA GCGGAGGCCTTGAGGTTTTG XACGGAGACTCT 2380 AACAGAAACTCT TATGGAGACTCT TATGGAGACTCT TATGGAGACTTT -ACGGAGACTCT TATGGAGACTCT TATGGAGACTTT -ACGGAGACTT	AAGTAGAGACT AAGTAGAGACT AAGTAGAGACT AAGTAGAGAAA CTTGXAAAAGT CTTGCGAAAAAT CTAGAAAAAGT XACTGGTTTAA 2390 24 AGCTGGTTTAA GACTGGTTTAA GACTGGTTTAA TTTTGGCAGAA TTTTGGCAGAA TCATGGCAGAA	G phtA.SEQ G phtA.SEQ T phtD.SEQ - phtE.SEQ C Hajority ISO A phtA.SEQ C phtB.Seq A phtD.SEQ C phtE.SEQ A phtD.SEQ A phtB.SeQ
2163 G G A A C C A G A A G J 2209 G A A G C T G A A G A T 1351 G C T G A A G A T 2310 2234 A A A A A A G T T G A A G 2213 A A A A G G T T G A A G 2258 C T G T T A T T A A C G 1365	2320 CCCAACTCAAAGA AAAACTXAXAGAX 2320 CCCAACTCAAAGAA AAAACTGAGAGAG CTAAGATAGCAGAT ATXAAAXCCAATGC 2370 CTGAAGCCAATGC ATTAGACAAAATGC ATTAGACAAAATGC TCTTGGAACXAAGG TCTTCAAATTATGG ATTTGGCACCCAGG TCTTGGAACGAAAG	CAGAAGAACCTC CTGAAATTCCTC CTGAAATTCCTC 2330 GCAGAAGTTTTG GCTGAAGATTTA GCGGAGGCCTTGAGGTTTTG XACGGAGACTCT 2380 AACAGAAACTCT CAAAGAGACTCT TATGGAGACATT -ACGGAGACCTCT TATGGAGACATT -ACGGAGACCTCT TATGGAGACATT -ACGGAGACATT -ACGGAGACATT -ACGGAGACCTA 2430 ATAACAATAGTA ACAACAATAGTA	### CTT G C A G A CT A CT G C A CT CT CT CT CT CT CT A CT	G phtA.SEQ G phtB.seq T phtD.SEQ - phtE.SEQ C Hajority ISO A phtA.SEQ C phtB.seq A phtD.SEQ C phtE.SEQ A phtD.SEQ C phtE.SEQ A phtD.SEQ C phtB.seq A phtB.seq A phtB.seq A phtB.seq A phtB.seq C phtA.SEQ C phtB.seq

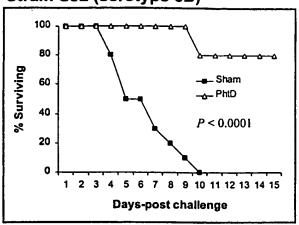
15/17

Figure 7(h)

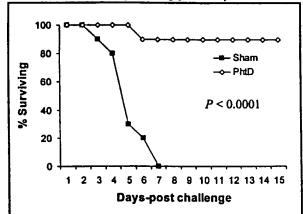
	<u>c</u>	λ	. 0		1	١,	<u> </u>	G	λ	C	7	,	. 7	. 7	. 0	3 (3 C	: 1	1	T	G	Ţ	7	A	λ	٨	a	G 2	L C	<u> </u>	G	T	λ	A 3	(]	_	λ	٨	G (. 1			_	-	_	c	Ŧ ·	<u>.</u> 1	taj
										2	4	50								2	47	10							:	24	80							2	190								25	00	
384	C	λ	. 0		V.	۱.،	N	À	X	Ţ	Ŧ	,	, C	7	-	. 0	3 C	G	1	Ŧ	G	T	ī	λ	A	A	λ	G	3 2		. 0	T	λ	A 1	· c	-	-	-		-	_	_	_	_	-	₹	7		ht
363	C	I	C	; ;	١.		Α.	A	X	C	T	,	1	. 1	: 0	; C	C	T	Ţ	T	λ	Ŧ	T	Y	Y	λ	G	G 1	V G	,	•	T	A.	A C	T	A	A	A	GC	T	, y	. Q	A	A	G	C	T		h
408	_						•	0	T	C	Ţ	C	7	7	: 0	: 0	C	T	I	T	G	T	T	Y	λ	λ	Y	G 2	r y		G	Ŧ	-		-	•	-	-		-	-	-	-	-	-	-	- (: r	ht
409	•	λ	. 0			. (3	G	X	C	Ţ	-	•	7	. 0	, ,		-	-	-	-	-	-	-	-	X	G.	Y 1	/ C	Y	G	C.	A.	Y Y	T	T	λ	X.	G (T		-	-	-	-	-		- 1	ht
																							_		_	_	_																	•••					
	λ	<u> </u>	_			_	-	<u>-</u>	=	=	-		- 1		. 0	<u> </u>	<u>: -</u>	<u> </u>	C	T	÷	=	0	_	Ç	۸	λ		<u>, y</u>	-	<u>.</u>	-	T		<u> </u>	<u>, y</u>	T	G	<u> </u>	<u> </u>	. λ	<u>. </u>		G	<u>c</u>	T	X A	<u>,</u> ,	ia :
										2	5	10								2	52	0							1	253	10							2:	40								25	50	
123	C	λ	-	•			•	-	-	•	T	C	7	G	_	-	-	_	-	T	λ	A	G	-	•	•	-			-	-	-	-		T	A	λ	6	G J		Ā	À	λ	λ	Ŧ	-		-	ht
113	Å	A	a	•	3 (•	:	3	X	A	T	1	1	G	0	1 0	: A	, C	C	C	X	G	G	λ	C	X	X.	c ı	l A	. T	A	C	T.	l I	T	A	T	G	3 C		G	λ	A	G	C	T (G ,	, ,	ht
141	λ	¥	-	•	• •	•	•	-	•	-	-	C	C	0	1 0	1 0	: T	C	C	T	A	-	-	-	-	-	-		• -	-	-	-	-		Ŧ	A	T	X I	3 1	À		A	A	Ç	C	T '	T J	L p	ht
140	-	-	-	•	• (1 (2 (3	•	-	-	-	-	-		•		-	-	-	-	-	-	-	C	X	λ.	X J	١,	C	λ	Ŧ	T		-	-	T	λ	3										ht
	λ	<u> </u>	λ			: 7	٠.	<u> </u>	x	X	•																																					ŀ	(a)
445	-	<u>,</u>	_	;		. ,	۲.	<u>.</u>	À	_	•																																						phi
463	À	A	A	. 1	. (: 1		A.	T	T																																						-	, ph
68	A	G	-	•			•	-	C	C																												•											phi
155																																																2	pht

Figure 8

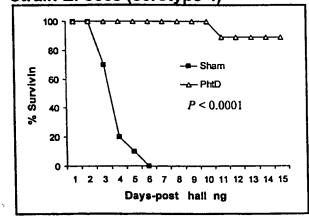
A. Strain SJ2 (serotype 6B)



B. Strain EF6796 (serotype 6A)



C. Strain EF5668 (serotype 4)



17/17

SEQUENCE LISTING

```
<110> Johnson, Leslie S.
      Koenig, Scott
     Adamou, John E.
<120> Streptococcus Pneumoniae and Immunogenic Fragments for
<130> 469201-444
<140>
<141>
<150> 60/113,048
<151> 1998-12-21
<160> 11
<170> PatentIn Ver. 2.0
<210> 1
<211> 36
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:Forward primer
      used in amplification of the Sp36 gene sequence.
<400> 1
atcggatcct tcttacgagt tgggactgta tcaagc
                                                                   36
<210> 2
<211> 35
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Forward primer
      used in amplification of the Sp36 gene sequence.
<400> 2
atcggatcca ctgtatcaag ctagaacggt taagg
                                                                   35
<210> 3
<211> 40
<212> DNA
```

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Reverse primer used in amplification of the Sp36 gene sequence.

<400> 3

agtcaagctt gtttattttt tccttactta cagatgaagg

40

- <210> 4
- <211> 838
- <212> PRT
- <213> Streptococcus pneumoniae

<400> 4

Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Val Leu Ala 1 5 10 15

Leu Ser Val Cys Ser Tyr Glu Leu Gly Arg His Gln Ala Gly Gln Val 20 25 30

Lys Lys Glu Ser Asn Arg Val Ser Tyr Ile Asp Gly Asp Gln Ala Gly
35 40 45

Gln Lys Ala Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly
50 55 60

Ile Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val 65 70 75 80

Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr
85 90 95

Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln
100 105 110

Leu Lys Asp Ser Asp Ile Val Asn Glu Ile Lys Gly Gly Tyr Val Ile 115 120 125

Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140

Asp Asn Ile Arg Thr Lys Glu Glu Ile Lys Arg Gln Lys Gln Glu His 145 150 155 160

Ser His Asn His Gly Gly Gly Ser Asn Asp Gln Ala Val Val Ala Ala 165 170 175

Arg Ala Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn Ala 180 185 190

- Ser Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His Gly
 195 200 205
- Asp His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu Leu 210 215 220
- Ala Ala Ala Glu Ala Tyr Trp Asn Gly Lys Gln Gly Ser Arg Pro Ser 225 230 235 240
- Ser Ser Ser Tyr Asn Ala Asn Pro Ala Gln Pro Arg Leu Ser Glu 245 250 255
- Asn His Asn Leu Thr Val Thr Pro Thr Tyr His Gln Asn Gln Gly Glu 260 265 270
- Asn Ile Ser Ser Leu Leu Arg Glu Leu Tyr Ala Lys Pro Leu Ser Glu 275 280 285
- Arg His Val Glu Ser Asp Gly Leu Ile Phe Asp Pro Ala Gln Ile Thr 290 295 300
- Ser Arg Thr Ala Arg Gly Val Ala Val Pro His Gly Asn His Tyr His 305 310 315 320
- Phe Ile Pro Tyr Glu Gln Met Ser Glu Leu Glu Lys Arg Ile Ala Arg
 325 330 335
- Ile Ile Pro Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg 340 345 350
- Pro Glu Gln Pro Ser Pro Gln Ser Thr Pro Glu Pro Ser Pro Ser Pro 355 360 365
- Gln Pro Ala Pro Asn Pro Gln Pro Ala Pro Ser Asn Pro Ile Asp Glu 370 375 380
- Lys Leu Val Lys Glu Ala Val Arg Lys Val Gly Asp Gly Tyr Val Phe 385 390 395 400
- Glu Glu Asn Gly Val Ser Arg Tyr Ile Pro Ala Lys Asp Leu Ser Ala 405 410 415
- Glu Thr Ala Ala Gly Ile Asp Ser Lys Leu Ala Lys Gln Glu Ser Leu
 420 425 430

Ser His Lys Leu Gly Ala Lys Lys Thr Asp Leu Pro Ser Ser Asp Arg 435 440 445

- Glu Phe Tyr Asn Lys Ala Tyr Asp Leu Leu Ala Arg Ile His Gln Asp 450 455 460
- Leu Leu Asp Asn Lys Gly Arg Gln Val Asp Phe Glu Ala Leu Asp Asn 465 470 475 480
- Leu Leu Glu Arg Leu Lys Asp Val Pro Ser Asp Lys Val Lys Leu Val
 485 490 495
- Asp Asp Ile Leu Ala Phe Leu Ala Pro Ile Arg His Pro Glu Arg Leu 500 505 510
- Gly Lys Pro Asn Ala Gln Ile Thr Tyr Thr Asp Asp Glu Ile Gln Val 515 520 525
- Ala Lys Leu Ala Gly Lys Tyr Thr Thr Glu Asp Gly Tyr Ile Phe Asp 530 535 540
- Pro Arg Asp Ile Thr Ser Asp Glu Gly Asp Ala Tyr Val Thr Pro His 545 550 550 560
- Met Thr His Ser His Trp Ile Lys Lys Asp Ser Leu Ser Glu Ala Glu 565 570 575
- Arg Ala Ala Gln Ala Tyr Ala Lys Glu Lys Gly Leu Thr Pro Pro 580 585 590
- Ser Thr Asp His Gln Asp Ser Gly Asn Thr Glu Ala Lys Gly Ala Glu
 595 600 605
- Ala Ile Tyr Asn Arg Val Lys Ala Ala Lys Lys Val Pro Leu Asp Arg 610 615 620
- Met Pro Tyr Asn Leu Gln Tyr Thr Val Glu Val Lys Asn Gly Ser Leu 625 630 635 640
- Ile Ile Pro His Tyr Asp His Tyr His Asn Ile Lys Phe Glu Trp Phe 645 650 655
- Asp Glu Gly Leu Tyr Glu Ala Pro Lys Gly Tyr Thr Leu Glu Asp Leu 660 665 670
- Leu Ala Thr Val Lys Tyr Tyr Val Glu His Pro Asn Glu Arg Pro His 675 680 685

Ser Asp Asn Gly Phe Gly Asn Ala Ser Asp His Val Arg Lys Asn Lys 690 695 700 Val Asp Gln Asp Ser Lys Pro Asp Glu Asp Lys Glu His Asp Glu Val 710 715 Ser Glu Pro Thr His Pro Glu Ser Asp Glu Lys Glu Asn His Ala Gly 725 730 Leu Asn Pro Ser Ala Asp Asn Leu Tyr Lys Pro Ser Thr Asp Thr Glu 745 Glu Thr Glu Glu Glu Ala Glu Asp Thr Thr Asp Glu Ala Glu Ile Pro 755 760 765 Gln Val Glu Asn Ser Val Ile Asn Ala Lys Ile Ala Asp Ala Glu Ala 770 775 780 Leu Leu Glu Lys Val Thr Asp Pro Ser Ile Arg Gln Asn Ala Met Glu 785 790 795 Thr Leu Thr Gly Leu Lys Ser Ser Leu Leu Leu Gly Thr Lys Asp Asn 805 Asn Thr Ile Ser Ala Glu Val Asp Ser Leu Leu Ala Leu Leu Lys Glu 825 Ser Gln Pro Ala Pro Ile 835 <210> 5 <211> 2531 <212> DNA <213> Streptococcus pneumoniae atgaaaatta ataaaaaata tctagcaggt tcagtggcag tccttgccct aagtgtttgt 60 tectatgaac ttggtegtea ecaagetggt caggttaaga aagagtetaa tegagtttet 120 tatatagatg gtgatcaggc tggtcaaaag gcagaaaact tgacaccaga tgaagtcagt 180 aagagggagg ggatcaacgc cgaacaaatc gtcatcaaga ttacggatca aggttatgtg 240 acctctcatg gagaccatta tcattactat aatggcaagg tcccttatga tgccatcatc 300 agtgaagagc tcctcatgaa agatccgaat tatcagttga aggattcaga cattgtcaat 360 gaaatcaagg gtggttatgt tatcaaggta gatggaaaat actatgttta ccttaaggat 420 gcagctcatg cggataatat tcggacaaaa gaagagatta aacgtcagaa gcaggaacac 480 agtcataatc acgggggtgg ttctaacgat caagcagtag ttgcagccag agcccaagga 540 cgctatacaa cggatgatgg ttatatcttc aatgcatctg atatcattga ggacacgggt 600

gatgettata tegtteetea eggegaecat taccattaca tteetaagaa tgagttatea 660

PCT/US99/30390 WO 00/37105

```
gctagcgagt tagctgctgc agaagcctat tggaatggga agcagggatc tcgtccttct 720
tcaagttcta gttataatgc aaatccagct caaccaagat tgtcagagaa ccacaatctg 780
actgtcactc caacttatca tcaaaatcaa ggggaaaaca tttcaagcct tttacgtgaa 840
ttgtatgcta aaccettate agaacgccat gtggaatetg atggcettat tttcgaccca 900
gcgcaaatca caagtcgaac cgccagaggt gtagctgtcc ctcatggtaa ccattaccac 960
tttatccctt atgaacaaat gtctgaattg gaaaaacgaa ttgctcgtat tattccctt 1020
cgttatcgtt caaaccattg ggtaccagat tcaagaccag aacaaccaag tccacaatcg 1080
actocggaac ctagtocaag teegcaacet geaccaaate eteaaceage teeaageaat 1140
ccaattgatg agaaattggt caaagaagct gttcgaaaag taggcgatgg ttatgtcttt 1200
gaggagaatg gagtttctcg ttatatccca gccaaggatc tttcagcaga aacagcagca 1260
ggcattgata gcaaactggc caagcaggaa agtttatctc ataagctagg agctaagaaa 1320
actgacctcc catctagtga tcgagaattt tacaataagg cttatgactt actagcaaga 1380
attcaccaag atttacttga taataaaggt cgacaagttg attttgaggc tttggataac 1440
ctgttggaac gactcaagga tgtcccaagt gataaagtca agttagtgga tgatattctt 1500
gccttcttag ctccgattcg tcatccagaa cgtttaggaa aaccaaatgc gcaaattacc 1560
tacactgatg atgagattca agtagccaag ttggcaggca agtacacaac agaagacggt 1620
tatatetttg atectegtga tataaceagt gatgaggggg atgeetatgt aacteeacat 1680
atgacccata gccactggat taaaaaagat agtttgtctg aagctgagag agcggcagcc 1740
caggettatg ctaaagagaa aggtttgacc ceteettega cagaceatea ggatteagga 1800
aatactgagg caaaaggagc agaagctatc tacaaccgcg tgaaagcagc taagaaggtg 1860
ccacttgatc gtatgcctta caatcttcaa tatactgtag aagtcaaaaa cggtagttta 1920
atcatacctc attatgacca ttaccataac atcaaatttg agtggtttga cgaaggcctt 1980
tatgaggcac ctaaggggta tactcttgag gatcttttgg cgactgtcaa gtactatgtc 2040
gaacatccaa acgaacgtcc gcattcagat aatggttttg gtaacgctag cgaccatgtt 2100
cgtaaaaata aggtagacca agacagtaaa cctgatgaag ataaggaaca tgatgaagta 2160
agtgagccaa ctcaccctga atctgatgaa aaagagaatc acgctggttt aaatccttca 2220
gcagataatc tttataaacc aagcactgat acggaagaga cagaggaaga agctgaagat 2280
accacagatg aggetgaaat teeteaagta gagaattetg ttattaacge taagatagea 2340
gatgcggagg ccttgctaga aaaagtaaca gatcctagta ttagacaaaa tgctatggag 2400
acattgactg gtctaaaaag tagtcttctt ctcggaacga aagataataa cactatttca 2460
gcagaagtag atagtctctt ggctttgtta aaagaaagtc aaccggctcc tatatagtaa 2520
aagcttaagc c
                                                                  2531
<210> 6
<211> 484
```

```
<212> PRT
<213> Streptococcus pneumoniae
```

<400> 6 Met Lys Phe Ser Lys Lys Tyr Ile Ala Ala Gly Ser Ala Val Ile Val 1 5 10 15

Ser Leu Ser Leu Cys Ala Tyr Ala Leu Asn Gln His Arg Ser Gln Glu 25 30

Asn Lys Asp Asn Asn Arg Val Ser Tyr Val Asp Gly Ser Gln Ser Ser 40

Gln Lys Ser Glu Asn Leu Thr Pro Asp Gln Val Ser Gln Lys Glu Gly 50 55 60

- Ile Gln Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val 65 70 75 80
- Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr
 85 90 95
- Asp Ala Leu Phe Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln
 100 105 110
- Leu Lys Asp Ala Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Ile Ile
 115 120 125
- Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140
- Asp Asn Val Arg Thr Lys Asp Glu Ile Asn Arg Gln Lys Gln Glu His 145 150 155 160
- Val Lys Asp Asn Glu Lys Val Asn Ser Asn Val Ala Val Ala Arg Ser 165 170 175
- Gln Gly Arg Tyr Thr Thr Asn Asp Gly Tyr Val Phe Asn Pro Ala Asp 180 185 190
- Ile Ile Glu Asp Thr Gly Asn Ala Tyr Ile Val Pro His Gly Gly His
 195 200 205
- Tyr His Tyr Ile Pro Lys Ser Asp Leu Ser Ala Ser Glu Leu Ala Ala 210 215 220
- Ala Lys Ala His Leu Ala Gly Lys Asn Met Gln Pro Ser Gln Leu Ser 225 230 235 240
- Tyr Ser Ser Thr Ala Ser Asp Asn Asn Thr Gln Ser Val Ala Lys Gly
 245 250 255
- Ser Thr Ser Lys Pro Ala Asn Lys Ser Glu Asn Leu Gln Ser Leu Leu 260 265 270
- Lys Glu Leu Tyr Asp Ser Pro Ser Ala Gln Arg Tyr Ser Glu Ser Asp 275 280 285
- Gly Leu Val Phe Asp Pro Ala Lys Ile Ile Ser Arg Thr Pro Asn Gly 290 295 300

Val Ala Ile Pro His Gly Asp His Tyr His Phe Ile Pro Tyr Ser Lys 305 310 315 320

- Leu Ser Ala Leu Glu Glu Lys Ile Ala Arg Met Val Pro Ile Ser Gly
 325 330 335
- Thr Gly Ser Thr Val Ser Thr Asn Ala Lys Pro Asn Glu Val Val Ser 340 345 350
- Ser Leu Gly Ser Leu Ser Ser Asn Pro Ser Ser Leu Thr Thr Ser Lys 355 360 365
- Glu Leu Ser Ser Ala Ser Asp Gly Tyr Ile Phe Asn Pro Lys Asp Ile 370 375 380
- Val Glu Glu Thr Ala Thr Ala Tyr Ile Val Arg His Gly Asp His Phe 385 390 395 400
- His Tyr Ile Pro Lys Ser Asn Gln Ile Gly Gln Pro Thr Leu Pro Asn 405 410 415
- Asn Ser Leu Ala Thr Pro Ser Pro Ser Leu Pro Ile Asn Pro Gly Thr
 420 425 430
- Ser His Glu Lys His Glu Glu Asp Gly Tyr Gly Phe Asp Ala Asn Arg
 435
 440
 445
- Ile Ile Ala Glu Asp Glu Ser Gly Phe Val Met Ser His Gly Asp His
 450 455 460
- Asn His Tyr Phe Phe Lys Lys Asp Leu Thr Glu Glu Gln Ile Lys Val 465 470 475 480

Arg Lys Asn Ile

<210> 7

<211> 1455

<212> DNA

<213> Streptococcus pneumoniae

<400> 7

atgaaattta gtaaaaata tatagcagct ggatcagctg ttatcgtatc cttgagtcta 60 tgtggcctatg cactaaacca gcatcgttcg caggaaaata aggacaataa tcgtgtctct 120 tatgtggatg gcagccagtc aagtcagaaa agtgaaaact tgacaccaga ccaggttagc 180 cagaaagaag gaattcaggc tgagcaaatt gtaatcaaaa ttacagatca gggctatgta 240 acgtcaccag gtgaccacta tcattactat aatgggaaag ttccttatga tgccctcttt 300

```
agtgaagaac tettgatgaa ggatecaaac tateaaetta aagaegetga tattgteaat 360
gaagtcaagg gtggttatat catcaaggtc gatggaaaat attatgtcta cctgaaagat 420
gcagctcatg ctgataatgt tcgaactaaa gatgaaatca atcgtcaaaa acaagaacat 480
gtcaaagata atgagaaggt taactctaat gttgctgtag caaggtctca gggacgatat 540
acgacaaatg atggttatgt ctttaatcca gctgatatta tcgaagatac gggtaatgct 600
tatatcgttc ctcatggagg tcactatcac tacattccca aaagcgattt atctgctagt 660
gaattagcag cagctaaagc acatctggct ggaaaaaata tgcaaccgag tcagttaagc 720
tattetteaa cagetagtga caataacaeg caatetgtag caaaaggate aactageaag 780
ccagcaaata aatctgaaaa tctccagagt cttttgaagg aactctatga ttcacctagc 840
gcccaacgtt acagtgaatc agatggcctg gtctttgacc ctgctaagat tatcagtcgt 900
acaccaaatg gagttgcgat tccgcatggc gaccattacc actttattcc ttacagcaag 960
ctttctgcct tagaagaaaa gattgccaga atggtgccta tcagtggaac tggttctaca 1020
gtttctacaa atgcaaaacc taatgaagta gtgtctagtc taggcagtct ttcaagcaat 1080
ccttcttctt taacgacaag taaggagete tetteageat etgatggtta tatttttaat 1140
ccaaaagata tcgttgaaga aacggctaca gcttatattg taagacatgg tgatcatttc 1200
cattacattc caaaatcaaa tcaaattggg caaccgactc ttccaaacaa tagtctagca 1260
acaccttctc catctcttcc aatcaatcca ggaacttcac atgagaaaca tgaagaagat 1320
ggatacggat ttgatgctaa tcgtattatc gctgaagatg aatcaggttt tgtcatgagt 1380
cacggagacc acaatcatta tttcttcaag aaggacttga cagaagagca aattaaggtg 1440
cgcaaaaaca tttag
<210> 8
<211> 819
<212> PRT
<213> Streptococcus pneumoniae
<400> 8
Met Lys Ile Asn Lys Lys Tyr Leu Val Gly Ser Ala Ala Ala Leu Ile
                  5
                                     10
Leu Ser Val Cys Ser Tyr Glu Leu Gly Leu Tyr Gln Ala Arg Thr Val
             20
                                 25
Lys Glu Asn Asn Arg Val Ser Tyr Ile Asp Gly Lys Gln Ala Thr Gln
                             40
```

Lys Thr Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly Ile
50 55 60

Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val Thr
65 70 75 80

Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr Asp 85 90 95

Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Lys Leu 100 105 110

Lys Asp Glu Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Val Ile Lys
115 120 125

- Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala Asp 130 135 140
- Asn Val Arg Thr Lys Glu Glu Ile Asn Arg Gln Lys Gln Glu His Ser 145 150 155 160
- Gln His Arg Glu Gly Gly Thr Pro Arg Asn Asp Gly Ala Val Ala Leu 165 170 175
- Ala Arg Ser Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn 180 185 190
- Ala Ser Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His 195 200 205
- Gly Asp His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu 210 215 220
- Leu Ala Ala Glu Ala Phe Leu Ser Gly Arg Gly Asn Leu Ser Asn 225 230 235 240
- Ser Arg Thr Tyr Arg Arg Gln Asn Ser Asp Asn Thr Ser Arg Thr Asn 245 250 255
- Trp Val Pro Ser Val Ser Asn Pro Gly Thr Thr Asn Thr Asn Thr Ser 260 265 270
- Asn Asn Ser Asn Thr Asn Ser Gln Ala Ser Gln Ser Asn Asp Ile Asp 275 280 285
- Ser Leu Leu Lys Gln Leu Tyr Lys Leu Pro Leu Ser Gln Arg His Val 290 295 300
- Glu Ser Asp Gly Leu Val Phe Asp Pro Ala Gln Ile Thr Ser Arg Thr 305 310 310 315 320
- Ala Arg Gly Val Ala Val Pro His Gly Asp His Tyr His Phe Ile Pro 325 330 335
- Tyr Ser Gln Met Ser Glu Leu Glu Glu Arg Ile Ala Arg Ile Ile Pro 340 345 350
- Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg Pro Glu Gln 355 360 365

Pro Ser Pro Gln Pro Thr Pro Glu Pro Ser Pro Gly Pro Gln Pro Ala 370 375 380

- Pro Asn Leu Lys Ile Asp Ser Asn Ser Ser Leu Val Ser Gln Leu Val
 385 390 395 400
- Arg Lys Val Gly Glu Gly Tyr Val Phe Glu Glu Lys Gly Ile Ser Arg 405 410 415
- Tyr Val Phe Ala Lys Asp Leu Pro Ser Glu Thr Val Lys Asn Leu Glu
 420 425 430
- Ser Lys Leu Ser Lys Gln Glu Ser Val Ser His Thr Leu Thr Ala Lys 435 440 445
- Lys Glu Asn Val Ala Pro Arg Asp Gln Glu Phe Tyr Asp Lys Ala Tyr 450 455 460
- Asn Leu Leu Thr Glu Ala His Lys Ala Leu Phe Glu Asn Lys Gly Arg 465 470 475 480
- Asn Ser Asp Phe Gln Ala Leu Asp Lys Leu Leu Glu Arg Leu Asn Asp 485 490 495
- Glu Ser Thr Asn Lys Glu Lys Leu Val Asp Asp Leu Leu Ala Phe Leu 500 505 510
- Ala Pro Ile Thr His Pro Glu Arg Leu Gly Lys Pro Asn Ser Gln Ile 515 520 525
- Glu Tyr Thr Glu Asp Glu Val Arg Ile Ala Gln Leu Ala Asp Lys Tyr 530 535 540
- Thr Thr Ser Asp Gly Tyr Ile Phe Asp Glu His Asp Ile Ile Ser Asp 545 550 555 560
- Glu Gly Asp Ala Tyr Val Thr Pro His Met Gly His Ser His Trp Ile
 565 570 575
- Gly Lys Asp Ser Leu Ser Asp Lys Glu Lys Val Ala Ala Gln Ala Tyr
 580 585 590
- Thr Lys Glu Lys Gly Ile Leu Pro Pro Ser Pro Asp Ala Asp Val Lys
 595 600 605
- Ala Asn Pro Thr Gly Asp Ser Ala Ala Ala Ile Tyr Asn Arg Val Lys 610 615 620

Gly Glu Lys Arg Ile Pro Leu Val Arg Leu Pro Tyr Met Val Glu His 625 630 635 640

Thr Val Glu Val Lys Asn Gly Asn Leu Ile Ile Pro His Lys Asp His 645 650 655

Tyr His Asn Ile Lys Phe Ala Trp Phe Asp Asp His Thr Tyr Lys Ala 660 665 670

Pro Asn Gly Tyr Thr Leu Glu Asp Leu Phe Ala Thr Ile Lys Tyr Tyr 675 680 685

Val Glu His Pro Asp Glu Arg Pro His Ser Asn Asp Gly Trp Gly Asn 690 695 700

Ala Ser Glu His Val Leu Gly Lys Lys Asp His Ser Glu Asp Pro Asn 705 710 715 720

Lys Asn Phe Lys Ala Asp Glu Glu Pro Val Glu Glu Thr Pro Ala Glu
725 730 735

Pro Glu Val Pro Gln Val Glu Thr Glu Lys Val Glu Ala Gln Leu Lys
740 745 750

Glu Ala Glu Val Leu Leu Ala Lys Val Thr Asp Ser Ser Leu Lys Ala 755 760 765

Asn Ala Thr Glu Thr Leu Ala Gly Leu Arg Asn Asn Leu Thr Leu Gln
770 780

Ile Met Asp Asn Asn Ser Ile Met Ala Glu Ala Glu Lys Leu Leu Ala
785 790 795 800

Leu Leu Lys Gly Ser Asn Pro Ser Ser Val Ser Lys Glu Lys Ile Asn 805 810 815

Lys Leu Asn

<210> 9

<211> 2451

<212> DNA

<213> Streptococcus pneumoniae

<400> 9

atgaaaatta ataagaaata ccttgttggt tctgcggcag ctttgatttt aagtgtttgt 60 tcttacgagt tgggactgta tcaagctaga acggttaagg aaaataatcg tgtttcctat 120

```
atagatggaa aacaagcgac gcaaaaaacg gagaatttga ctcctgatga ggttagcaag 180
cgtgaaggaa tcaatgctga gcaaatcgtc atcaagataa cagaccaagg ctatgtcact 240
tcacatggcg accactatca ttattacaat ggtaaggttc cttatgacgc tatcatcagt 300
gaagaattac tcatgaaaga tccaaactat aagctaaaag atgaggatat tgttaatgag 360
gtcaagggtg gatatgttat caaggtagat ggaaaatact atgtttacct taaggatgct 420
gcccacgcgg ataacgtccg tacaaaagag gaaatcaatc gacaaaaaca agagcatagt 480
caacatcgtg aaggtggaac tccaagaaac gatggtgctg ttgccttggc acgttcgcaa 540
ggacgctata ctacagatga tggttatatc tttaatgctt ctgatatcat agaggatact 600
ggtgatgctt atatcgttcc tcatggagat cattaccatt acattcctaa gaatgagtta 660
tcagctagcg agttggctgc tgcagaagcc ttcctatctg gtcgaggaaa tctgtcaaat 720
tcaagaacct atcgccgaca aaatagcgat aacacttcaa gaacaaactg ggtaccttct 780
gtaagcaatc caggaactac aaatactaac acaagcaaca acagcaacac taacagtcaa 840
gcaagtcaaa gtaatgacat tgatagtctc ttgaaacagc tctacaaact gcctttgagt 900
caacgacatg tagaatctga tggccttgtc tttgatccag cacaaatcac aagtcgaaca 960
gctagaggtg ttgcagtgcc acacggagat cattaccact tcatccctta ctctcaaatg 1020
tetgaattgg aagaacgaat egetegtatt atteeeette gttategtte aaaccattgg 1080
gtaccagatt caaggccaga acaaccaagt ccacaaccga ctccggaacc tagtccaggc 1140
ccgcaacctg caccaaatct taaaatagac tcaaattctt ctttggttag tcagctggta 1200
cgaaaagttg gggaaggata tgtattcgaa gaaaagggca tctctcgtta tgtctttgcg 1260
aaagatttac catctgaaac tgttaaaaat cttgaaagca agttatcaaa acaagagagt 1320
gtttcacaca ctttaactgc taaaaaagaa aatgttgctc ctcgtgacca agaattttat 1380
gataaagcat ataatctgtt aactgaggct cataaagcct tgtttgnaaa taagggtcgt 1440
aattotgatt tocaagoott agacaaatta ttagaacgot tgaatgatga atogactaat 1500
aaagaaaaat tggtagatga tttattggca ttcctagcac caattaccca tccagagcga 1560
cttggcaaac caaattctca aattgagtat actgaagacg aagttcgtat tgctcaatta 1620
gctgataagt atacaacgtc agatggttac atttttgatg aacatgatat aatcagtgat 1680
gaaggagatg catatgtaac gcctcatatg ggccatagtc actggattgg aaaagatagc 1740
ctttctgata aggaaaaagt tgcagctcaa gcctatacta aagaaaaagg tatcctacct 1800
ccatctccag acgcagatgt taaagcaaat ccaactggag atagtgcagc agctatttac 1860
aatcgtgtga aaggggaaaa acgaattcca ctcgttcgac ttccatatat ggttgagcat 1920
acagttgagg ttaaaaacgg taatttgatt attcctcata aggatcatta ccataatatt 1980
aaatttgett ggtttgatga teacacatae aaagetecaa atggetatae ettggaagat 2040
ttgtttgcga cgattaagta ctacgtagaa caccctgacg aacgtccaca ttctaatgat 2100
ggatggggca atgccagtga gcatgtgtta ggcaagaaag accacagtga agatccaaat 2160
aagaacttca aagcggatga agagccagta gaggaaacac ctgctgagcc agaagtccct 2220
caagtagaga ctgaaaaagt agaagcccaa ctcaaagaag cagaagtttt gcttgcgaaa 2280
gtaacggatt ctagtctgaa agccaatgca acagaaactc tagctggttt acgaaataat 2340
ttgactcttc aaattatgga taacaatagt atcatggcag aagcagaaaa attacttgcg 2400
ttgttaaaag gaagtaatcc ttcatctgta agtaaggaaa aaataaacta a
                                                                  2451
<210> 10
<211> 819
<212> PRT
<213> Streptococcus pneumoniae
<400> 10
Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Val Leu Ala
                                     10
```

Leu Ser Val Cys Ser Tyr Glu Leu Gly Arg Tyr Gln Ala Gly Gln Asp 20 25 30

- Lys Lys Glu Ser Asn Arg Val Ala Tyr Ile Asp Gly Asp Gln Ala Gly
 35 40 45
- Gln Lys Ala Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly
 50 55 60
- Ile Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val 65 70 75 80
- Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr 85 90 95
- Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln
 100 105 110
- Leu Lys Asp Ser Asp Ile Val Asn Glu Ile Lys Gly Gly Tyr Val Ile 115 120 125
- Lys Val Asn Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140
- Ser His Asn His Asn Ser Arg Ala Asp Asn Ala Val Ala Ala Ala Arg 165 170 175
- Ala Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn Ala Ser 180 185 190
- Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His Gly Asp 195 200 205
- His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu Leu Ala 210 215 220
- Ala Ala Glu Ala Tyr Trp Asn Gly Lys Gln Gly Ser Arg Pro Ser Ser 225 230 235 240
- Ser Ser Ser Tyr Asn Ala Asn Pro Ala Gln Pro Arg Leu Ser Glu Asn 245 250 255
- His Asn Leu Thr Val Thr Pro Thr Tyr His Gln Asn Gln Gly Glu Asn 260 265 270

Ile Ser Ser Leu Leu Arg Glu Leu Tyr Ala Lys Pro Leu Ser Glu Arg 275 280 285

- His Val Glu Ser Asp Gly Leu Ile Phe Asp Pro Ala Gln Ile Thr Ser 290 295 300
- Arg Thr Ala Arg Gly Val Ala Val Pro His Gly Asn His Tyr His Phe 305 310 315 320
- Ile Pro Tyr Glu Gln Met Ser Glu Leu Glu Lys Arg Ile Ala Arg Ile 325 330 335
- Ile Pro Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg Pro 340 345 350
- Glu Glu Pro Ser Pro Gln Pro Thr Pro Glu Pro Ser Pro Ser Pro Gln 355 360 365
- Pro Ala Pro Ser Asn Pro Ile Asp Gly Lys Leu Val Lys Glu Ala Val 370 375 380
- Arg Lys Val Gly Asp Gly Tyr Val Phe Glu Glu Asn Gly Val Ser Arg 385 390 395 400
- Tyr Ile Pro Ala Lys Asp Leu Ser Ala Glu Thr Ala Ala Gly Ile Asp 405 410 415
- Ser Lys Leu Ala Lys Gln Glu Ser Leu Ser His Lys Leu Gly Thr Lys 420 425 430
- Lys Thr Asp Leu Pro Ser Ser Asp Arg Glu Phe Tyr Asn Lys Ala Tyr 435 440 445
- Asp Leu Leu Ala Arg Ile His Gln Asp Leu Leu Asp Asn Lys Gly Arg 450 455 460
- Gln Val Asp Phe Glu Ala Leu Asp Asn Leu Leu Glu Arg Leu Lys Asp 465 470 475 480
- Val Ser Ser Asp Lys Val Lys Leu Val Glu Asp Ile Leu Ala Phe Leu 485 490 495
- Ala Pro Ile Arg His Pro Glu Arg Leu Gly Lys Pro Asn Ala Gln Ile 500 505 510
- Thr Tyr Thr Asp Asp Glu Ile Gln Val Ala Lys Leu Ala Gly Lys Tyr
 515 520 525

Thr Ala Glu Asp Gly Tyr Ile Phe Asp Pro Arg Asp Ile Thr Ser Asp 530 540

- Glu Gly Asp Ala Tyr Val Thr Pro His Met Thr His Ser His Trp Ile
 545 550 555 560
- Lys Lys Asp Ser Leu Ser Glu Ala Glu Arg Ala Ala Gln Ala Tyr 565 570 575
- Ala Glu Glu Lys Gly Leu Thr Pro Pro Ser Thr Asp His Gln Asp Ser 580 585 590
- Gly Asn Thr Glu Ala Lys Gly Ala Glu Ala Ile Tyr Asn Arg Val Lys
 595 600 605
- Ala Ala Lys Lys Val Pro Leu Asp Arg Met Pro Tyr Asn Leu Gln Tyr 610 615 620
- Thr Val Glu Val Lys Asn Gly Ser Leu Ile Ile Pro His Tyr Asp His 625 630 635 640
- Tyr His Asn Ile Lys Phe Glu Trp Phe Asp Glu Gly Leu Tyr Glu Ala 645 650 655
- Pro Lys Gly Tyr Thr Leu Glu Asp Leu Leu Ala Thr Val Lys Tyr Tyr 660 665 670
- Val Glu His Pro Asn Glu Arg Pro His Ser Asp Asn Gly Phe Gly Asn 675 680 685
- Ala Ser Asp His Val Gln Arg Asn Lys Asn Gly Gln Ala Asp Thr Asn 690 695 700
- Gln Thr Glu Lys Pro Ser Glu Glu Lys Pro Gln Thr Glu Lys Pro Glu 705 710 715 720
- Glu Glu Thr Pro Arg Glu Glu Lys Pro Gln Ser Glu Lys Pro Glu Ser
 725 730 735
- Pro Lys Pro Thr Glu Glu Pro Glu Glu Ser Pro Glu Glu Ser Glu Glu 740 745 750
- Pro Gln Val Glu Thr Glu Lys Val Glu Glu Lys Leu Arg Glu Ala Glu
 755 760 765
- Asp Leu Leu Gly Lys Ile Gln Asp Pro Ile Ile Lys Ser Asn Ala Lys
 770 775 780

Glu Thr Leu Thr Gly Leu Lys Asn Asn Leu Leu Phe Gly Thr Gln Asp
785 790 795 800

Asn Asn Thr Ile Met Ala Glu Ala Glu Lys Leu Leu Ala Leu Leu Lys 805 810 815

Glu Ser Lys

<210> 11

<211> 2531

<212> DNA

<213> Streptococcus pneumoniae

<400> 11

atgaaaatta ataaaaaata tctagcaggt tcagtggcag tccttgccct aagtgtttgt 60 teetatgage ttggaegtta ceaagetggt caggataaga aagagtetaa tegagttget 120 tatatagatg gtgatcaggc tggtcaaaag gcagaaaact tgacaccaga tgaagtcagt 180 aagagggagg ggatcaacgc cgaacaaatt gttatcaaga ttacggatca aggttatgtg 240 acctctcatg gagaccatta tcattactat aatggcaagg ttccttatga tgccatcatc 300 agtgaagagc tecteatgaa agateegaat tateagttga aggatteaga cattgteaat 360 gaaatcaagg gtggttatgt cattaaggta aacggtaaat actatgttta ccttaaggat 420 gcrgctcatg cggataatat tcggacaaaa gaagagatta aacgtcagaa gcaggaacgc 480 agtcataatc ataactcaag agcagataat gctgttgctg cagccagagc ccaaggacgt 540 tatacaacgg atgatgggta tatcttcaat gcatctgata tcattgagga cacgggtgat 600 gcttatatcg ttcctcacgg cgaccattac cattacattc ctaagaatga gttatcagct 660 agcgagttag ctgctgcaga agcctattgg aatgggaagc agggatctcg tccttcttca 720 agttctagtt ataatgcaaa tccagctcaa ccaagattgt cagagaacca caatctgact 780 gtcactccaa cttatcatca aaatcaaggg gaaaacattt caagcctttt acgtgaattg 840 tatgctaaac ccttatcaga acgccatgtg gaatctgatg gccttatttt cgacccagcg 900 caaatcacaa gtcgaaccgc cagaggtgta gctgtccctc atggtaacca ttaccacttt 960 atcccttatg aacaaatgtc tgaattggaa aaacgaattg ctcgtattat tccccttcgt 1020 tategtteaa accattgggt accagattea agaccagaag aaccaagtee acaaccgact 1080 ccagaaccta gtccaagtcc gcaaccagct ccaagcaatc caattgatgg gaaattggtc 1140 aaagaagetg ttcgaaaagt aggcgatggt tatgtctttg aggagaatgg agtttctcgt 1200 tatateceag ccaaggatet tteageagaa acageageag geattgatag caaactggee 1260 aagcaggaaa gtttatctca taagctagga actaagaaaa ctgacctccc atctagtgat 1320 cgagaatttt acaataaggc ttatgactta ctagcaagaa ttcaccaaga tttacttgat 1380 aataaaggtc gacaagttga ttttgaggct ttggataacc tgttggaacg actcaaggat 1440 gtctcaagtg ataaagtcaa gttagtggaa gatattcttg ccttcttagc tccgattcgt 1500 catccagaac gtttaggaaa accaaatgcg caaattacct acactgatga tgagattcaa 1560 gtagccaagt tggcaggcaa gtacacagca gaagacggtt atatetttga teetegtgat 1620 ataaccagtg atgaggggga tgcctatgta actccacata tgacccatag ccactggatt 1680 aaaaaagata gtttgtctga agctgagaga gcggcagccc aggcttatgc traagagaaa 1740 ggtttgaccc ctccttcgac agaccatcag gattcaggaa atactgaggc aaaaggagca 1800 gaagctatct acaaccgmgt gaaagcagct aagaaggtgc cacttgatcg tatgccttac 1860

aatetteaat atactgtaga agteaaaac ggtagtttaa teatacetea ttatgaceat 1920 taccataaca teaaatttga gtggtttgac gaaggeettt atgaggeace taaggggtat 1980 actettgagg atettttgge gactgteaag tactatgteg aacateeaaa egaacgteeg 2040 catteagata atggttttgg taacgetage gaccatgtte aaagaaacaa aaatggteaa 2100 getgatacea ateaaacgga aaaaccaage gaggagaaac eteagacaga aaaacetgag 2160 gaagaaacee etegagaaga gaaaceegeaa agegagaaac eagagtetee aaaaccaaca 2220 gaggaaacaa tgaggagge tgaagatta ettggaaaaa teeaggatee aattateaag 2340 teeaatgeea aagagactee eacaggatta aaaaataatt tactatttgg eaceecaggae 2400 aacaatacta ttatggeaga agetgaaaaa ettatggett tattaaagga gagtaagtaa 2460 aggtagaage ttaaagggega atttggeace eaggacaaca atactattat ggeagaaget 2520 gaaaaactat t